

caBench-to-Bedside 1.0

User's Manual

Document Change History

| Version Number | Date | Contributor | Description |
|----------------|---------------|--|-------------|
| V1.0 | July 12, 2007 | Washington University/Persistent Systems | Reviewed |



caBIG™ *cancer Biomedical Informatics Grid™*

an initiative of the National Cancer Institute

Updated 20 July

| Team(s) | | |
|---------------------------|-----------------------------|----------------------------------|
| <i>Development</i> | <i>User's Manual</i> | <i>Program Management</i> |
| Chandrakant Talele | Rahul Ner | Dr. Rakesh Nagarajan |
| Srinath K | Srinath K | Srikanth Adiga |
| Rahul Ner | Chandrakant Talele | |
| Deepak Shingan | Srikanth Adiga | |
| Chetan Patil | Raji Joshi | |
| Hrishikesh Rajpathak | Mukesh Sharma | |
| | | |

| <i>Contacts and Support</i> | |
|------------------------------------|--|
| Training contact | washu_cab2b@persistent.co.in |
| Customer support contact | http://cab2b.wustl.edu/cab2b/ |

Table of Contents

| | | |
|-----------|--|----|
| Chapter 1 | Introduction..... | 6 |
| | In this chapter | 6 |
| | Introduction to caB2B | 6 |
| | Getting started with caB2B | 7 |
| | Overview of the Manual | 7 |
| | Organization of the User Manual..... | 8 |
| | Document Text Conventions | 8 |
| Chapter 2 | Overview of the caB2B..... | 9 |
| | In this chapter | 9 |
| | Introduction to Category and Limits..... | 9 |
| | The Home page..... | 11 |
| | Search data – The Query Wizard..... | 11 |
| | Choose Search Category | 13 |
| | Define Limit..... | 16 |
| | Limit Set – Diagrammatic View | 17 |
| | Define Search Results View..... | 23 |
| | View Search Results | 25 |
| | Viewing and Saving Data List..... | 30 |
| Chapter 3 | Experiments..... | 33 |
| | In this Chapter | 33 |
| | What are Experiments and Projects?..... | 33 |
| | Creating an Experiment..... | 34 |
| | Viewing the Existing Experiments | 36 |
| | Opening an Experiment..... | 36 |
| | Working with an Experiment..... | 37 |

| | |
|--|----|
| Analyzing Data | 41 |
| Visualizing Data..... | 44 |
| Chapter 4 Appendix 1 | 49 |
| Examples of category search | 49 |
| Define Limit: Operators and Values | 53 |
| Record Details View | 55 |
| Administrator defined categories..... | 55 |
| BioDataCube | 56 |
| Types of Filters | 57 |

Chapter 1 Introduction

In this chapter

This chapter explains the need of caB2B application. It also helps you understand how to use caB2B to achieve those needs. It contains following sections:

- Introduction to caB2B
- Getting started with caB2B
- Overview of the manual
- Organization of the manual
- Document text conventions

Introduction to caB2B

The NCI caBIG™ project creates a common, extensible informatics platform that integrates diverse data types and supports interoperable analytic tools. caBIG™ is developing separate applications that will facilitate individual steps involved in micro-array analysis. These applications are also useful to bio-informaticians.

caGrid is the infrastructure for caBIG that helps integrate these applications. caGrid can be used to perform investigations involving the integration of data and analytical services from diverse research communities.

caBench-to-Bedside (caB2B) is an application that leverages these tools in a user-friendly

graphical user interface (GUI). caB2B can be used by the physician scientist to perform operations such as the following:

- **Query any caGrid data service to obtain data.**

The data service can be a single data service, multiple data services, or a combination of the two services using semantically interoperable Common Data Elements (CDEs).

For example, caB2B enables investigators to query tissue banks at multiple cancer centers, enabling studies to focus on very specific tumor subtypes and to target less common tumors by pooling bio-specimen resources. This helps in attenuating the problem of small sample sizes.

- **Collect data and create experiments.**

caB2B also enables investigators to perform novel in silico experiments using micro-array data sets that have already been collected. These capabilities facilitate the process of identifying genes that are up-regulated or down-regulated in specific cancers. These capabilities also enable investigators to view data in the context of biological pathways. Investigators also gain further understanding of the complex system of cancer biology by identifying genes important to the development and treatment of cancer. All of these lead to the more effective identification of novel drug targets, which improves treatment strategies.

- **Perform various analyses** by using various grid-enabled analytical services.
caB2B enables individual users to work on common platforms to access caGrid resources in real time from common data and analytical services.
- **Visualize analysis results** by using various viewers such as charts and dendrograms.

Getting started with caB2B

- To start caB2B client refer to Installation Guide.
- To search for the data sets refer to Search data – The Query Wizard
- To work on experiments refer to Experiments

Overview of the Manual

This manual provides the details of how to use caB2B application to

- Search the data sets of interests
- Create the experiment from the data sets
- Analyze those data
- Visualize

All the chapters in this user's manual begin with an introductory statement and followed by a list of topics included.





Organization of the User Manual

Following table summarizes the chapters in this manual.

| Chapter in caB2B | Chapter Contents |
|----------------------|--|
| 2. Overview of caB2B | General overview Searching categories Adding limits Defining output Viewing the results Saving data |
| 3. Experiments | Creating Experiments Analyzing data Visualizing data |
| 6. Appendix 1 | Examples of category search Operators in Defining limit Details of the various records Various filters |

Document Text Conventions

Following table illustrates documentation conventions represented in this manual.

| Convention | Description | Example | | |
|--|--|--|--|--|
|  | Indicates a information of particular interest | <table><tr><td></td><td>Administrator can create a new category.</td></tr></table> |  | Administrator can create a new category. |
|  | Administrator can create a new category. | | | |
| <div>CAUTION</div> | Indicates that you should be particularly aware of the information provided. | <table><tr><td>CAUTION</td><td>It deletes permanently.</td></tr></table> | CAUTION | It deletes permanently. |
| CAUTION | It deletes permanently. | | | |

Chapter 2 Overview of the caB2B

In this chapter

This chapter gives details of how to search data using caB2B. The following are the sections in this chapter:

- Introduction to some concepts
- Overview of Home Page
- Search data
- Choose Search Category
- Define Limit
- Define Search Results View
- View Search Results
- Viewing and Saving Data List

Introduction to Category and Limits

What is a Category?

A category is a collection of *searchable* attributes. You can use a category to search, view, and add data to experiments. Every UML class in an application is, trivially, a category. For example, Participant is a UML class from *caTissue Core* application.

Examples:

Participant is a category consisting of attributes like

- Identifier
- First Name
- Last Name
- Gender
- Race


Gene Annotation is a category consisting of attributes like

- Entrez Gene ID
- Gene Symbol
- Gene Name
- GenBank Accession Number

- Taxonomy ID
- Chromosome

Why do we need categories?

Categories are designed to help you to easily locate all the data elements needed for your search.


| | |
|---|--|
|  | <p>Categories are created by the administrator.</p> <p>Your administrator can create new categories based on your requirements. If there are queries that you perform regularly, discuss them with your administrator.</p> |
|---|--|

What is a Limit?

A limit is a condition (or criterion or predicate or condition) on a category. It is the smallest unit of a query. For example, *female participants*

What is a query?

The collection of limits is a query. For example, *female participants who have DNA specimens collected from their breast*. This query has three limits: *female participants*, *DNA specimens from breast*. There are two categories involved in this query: *Participant* and *Specimen*.

| | |
|---|---|
|  | <p>Your administrator can create a new category that contains relevant data elements from Specimens and Participants. That way you can enter all the limits in a single category.</p> |
|---|---|

The Home page

For details on installing and starting the caB2B client, refer to the *Client Installation* chapter of the Installation Manual.

When the caB2B application is launched, you see the caB2B Home page as shown below.

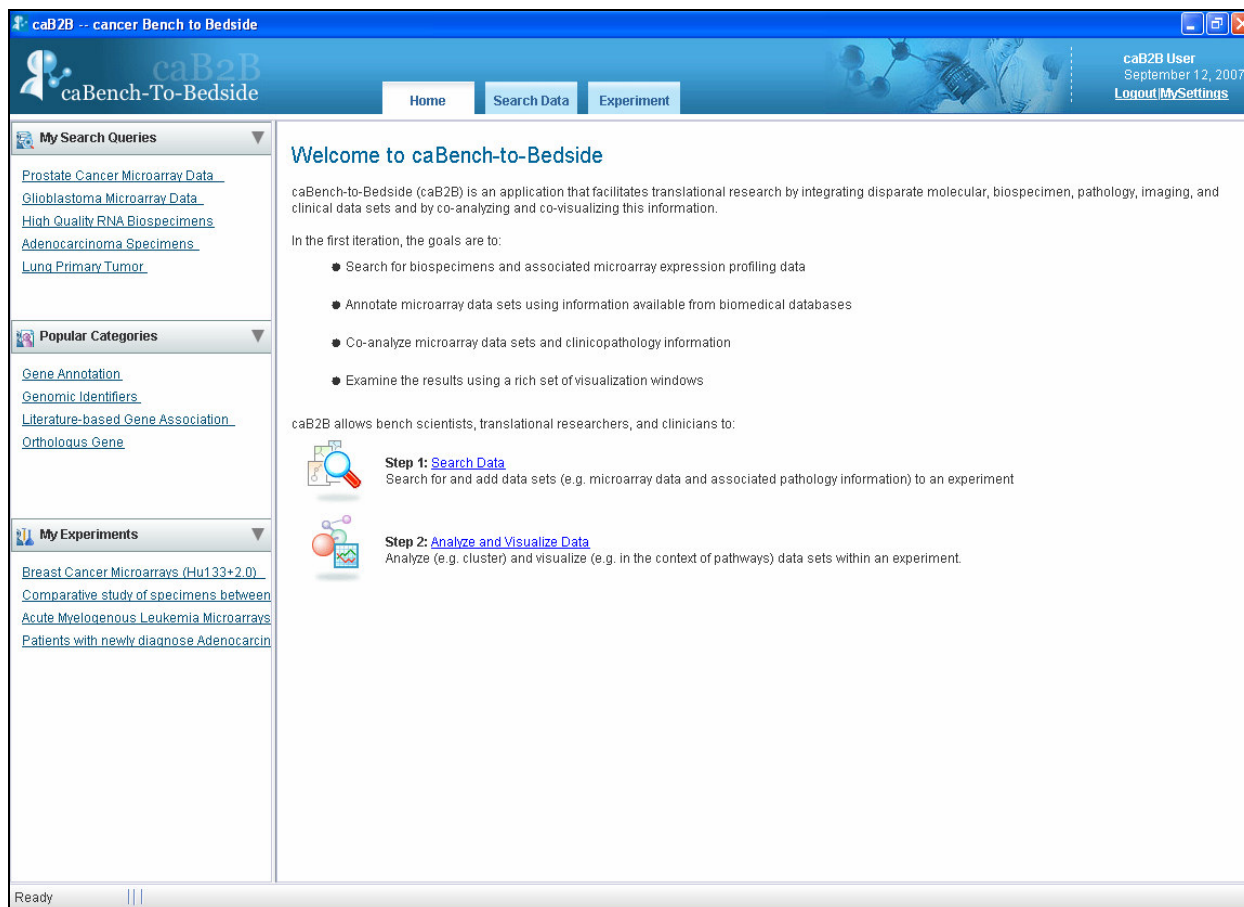


Figure 1: caB2B Home page

As mentioned in the introductory note, you would typically perform the following steps while using caB2B:

1. You would perform queries to obtain the data that you're interested in. Details of this step are provided in the [Searching Data](#) chapter.
2. You would then analyze and visualize the data you've gathered. Details of this step are provided in the [Experiment](#) chapter.

Search data – The Query Wizard

This section provides the details of the steps that you need to perform to obtain the data. Click the

Search Data link on the **Home** page to open the **Search data for Experiment** wizard (Figure 2: Choose Search Category Tab).

The wizard guides you through the following steps:

| STEP ONE | STEP TWO | STEP THREE | STEP FOUR | STEP FIVE |
|------------------------|------------|------------------------|---------------------|----------------|
| Choose Search Category | Add Limits | Choose Output Category | View Search Results | View Data List |

The steps used to search data are:

1. Choose Search Category: Search for Category on which you would like to apply limits on
2. Add Limits: Add limits on the Category
3. Choose Output Category (Optional): Select the output that you would like to see (i.e. define the results view)
4. View Search Results: View the results view in a Google-like search results screen and add data to the data list (or shopping cart).
5. View Data List: View the data you have added in the data list and create Experiment.

Choose Search Category

This step helps you search for the category on which you have limits.. For example, if you have a query with limits on participants, you can use this feature to first find the *Participant* category.

Click the category of your interest, to open a page to enter the limits on that category.

To find the categories, do the following:

1. Type the search string in the text box. *Gene, gender, rna, experiment* are examples of search strings.
2. Click the **Search** button or press ENTER.
3. Categories that match the string you entered are displayed as hyperlinks along with a short description.
4. Once you find a Category of your choice, click the **Category** to proceed to the next step.

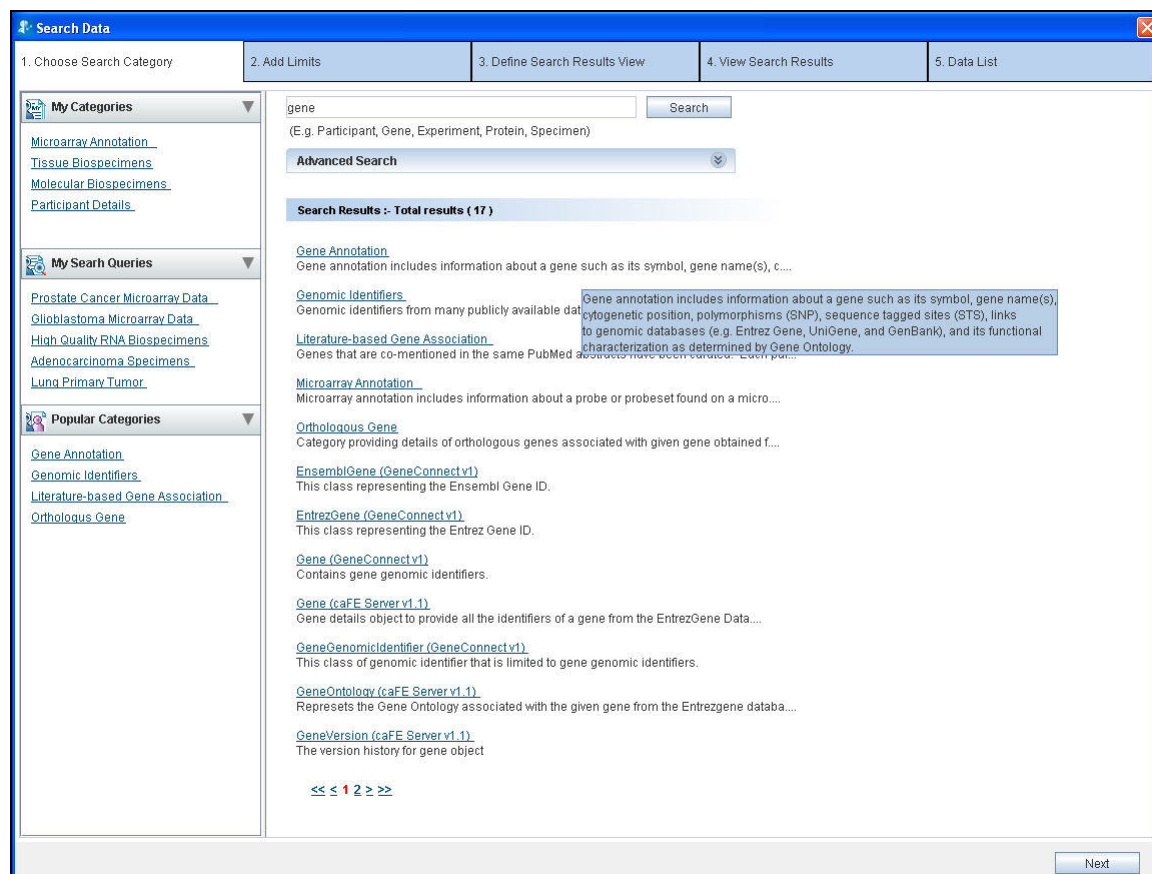


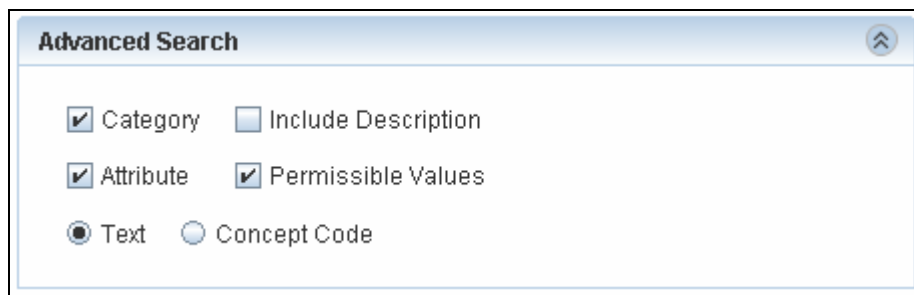
Figure 2: Choose Search Category Tab

More on searching:

- The search performed is partial and is not case-sensitive. Hence the search for terms such as *Gene*, *gen* or *GENE* returns the same results.
- If a category's description is long, you can view the complete description as a tool-tip by positioning the mouse over that description.
- If the number of matching categories is large, the system displays the results in multiple pages. Click the page numbers at the bottom to navigate to different pages.
- You can separate the multiple search terms using spaces or tabs. The system searches for each term individually and displays a collective set of matching categories. For example, if you enter the search string *Participant Gene*, you will find the categories that match either *Participant*, *Gene*, or both.

Advanced Search (*)

Expanding the **Advanced Search** box displays the advanced options that help you refine the category search. The following is the default selection of the various options:



The screenshot shows a window titled "Advanced Search" with a close button in the top right corner. Inside the window, there are several options with checkboxes or radio buttons:

- ☒ Category
- ☐ Include Description
- ☒ Attribute
- ☒ Permissible Values
- ☒ Text
- ☐ Concept Code

Figure 3: Advanced search

If the system returns too many results for your Category search, you can use the **Advanced Search** to fine-tune your search.

Table 3

| Option | Description |
|--------------------------|---|
| <i>Check Boxes</i> | |
| <i>Category</i> | Includes the category names in the search. |
| <i>Attribute</i> | <p>Attribute (also known as data element) is the lowest unit of the searchable items. A category contains many attributes.</p> <p>Select this check box to include the individual attribute names in the search.</p> <p>For example: select <i>attribute</i> and search for <i>gender</i>. You will find the category Participant since it contains an attribute by name <i>Gender</i>.</p> |
| <i>Permissible value</i> | <i>Some attributes contain a fixed set of values. For example, Gender can be Male, Female or Unspecified. Organ can be heart, lung, breast, and so on.</i> |

| <i>Option</i> | <i>Description</i> |
|----------------------------|---|
| | <p>This fixed set is known as Permissible Values.</p> <p>Select this check-box to include the permissible values in the search. So if you select the permissible value as <i>Male</i> the system displays the Participant category since it contains an attribute <i>Gender</i> which contains a permissible value <i>Male gender</i>.</p> |
| <i>Include Description</i> | Includes the description text in the search. |
| Option Buttons | |
| <i>Text</i> | <p>The system matches the search string that you type with the name of the category, its attributes, or the permissible values based on the text.</p> <p>This is the simple, most commonly used and the default way of searching a category.</p> |
| <i>Concept Code</i> | <p>The system treats the search string that you type as an EVS concept code. The system matches this concept code with the concept code of the category, its attributes, or the permissible values. The results contain matching categories .</p> <p>Note: Include Description check-box will be enabled only for a text based search when you select the Category check box, the Attribute check box, or both.</p> |

The various refinements that you can make using these options are illustrated in **Appendix 1**

Examples of category search.

Search Procedure

1. In the **Search** field, type the category you want to enter limits on.
2. Specify advanced options if required.
 - Select either the **Text** or the **Concept Code** option button.
 - Choose the required options from Category, Attribute, Include Description and Permissible Values.
3. Click on the **Search** button.
4. Matching categories are displayed. Each category has a hyperlink and contains a description associated with it.
5. If the size of the returned result set is large, the results are displayed in multiple pages.
6. Click on a category or click the **Next** button to proceed to the **Add limits** step.

Define Limit

In the previous step you have learned how to search for the category of your interest. Now you need to specify limits on that category so that the system can fetch the data.

A limit is a condition (or criterion or predicate) on a category and Multiple limits together form a query. The following table outlines two example queries and the limits added to them. This should give you an idea of how to create limits for your query.


Table 4

| Example | Category | Limits |
|---|---------------------------|---|
| Get the details about all the African male participants | Participant | Gender equals Male Race equals African |
| Biospecimens from prostate adenocarcinoma | Specimen Characteristics | Tissue Site equals prostate |
| | Specimen Collection Group | Clinical diagnosis equals adenocarcinoma |

Defining a Limit

To define a limit, you need to perform the following steps:

1. Choose an operator from the list of operators adjacent to the attribute you wish to constrain.
2. Specify the constraining values.

| | |
|---|---|
|  | For a detailed explanation of the various operators and how to specify values, refer to Define Limit: Operators and Values . |
|---|---|

Simple Example: Single constraint – Show all male participants

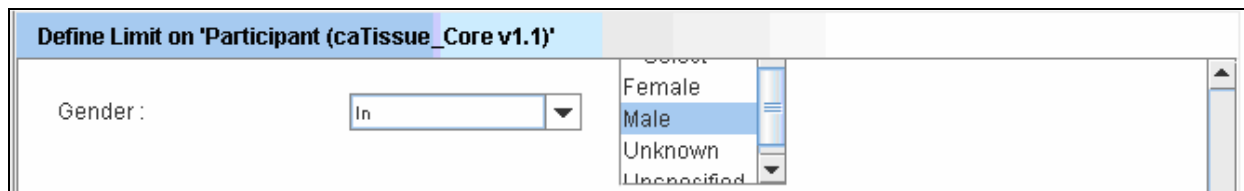


Figure 4: Male participants

Example with multiple constraints within one limit: male participants who are alive

Define Limit on 'Participant (caTissue_Core v1.1)'

Gender : In [Female, Male, Unknown, Unspecified]

Identifier : Between [] []

Last Name : Contains []

Middle Name : Contains []

Sex Genotype : In [-- Select --, Klinefelter's Syndrome, Mosaic including XXXXY, Not Specified, Karyo Y Syndrome]

Social Security Number : Contains []

Vital Status : In [-- Select --, Alive, Dead]

Figure 5: Male participants who are alive

CAUTION

Do not forget to click the **Add Limit** button at the bottom of the **Define Limits** screen before proceeding. If you forget to click this button, the constraints that you specify will not be saved.

Limit Set – Diagrammatic View

The **Limit Set** pane displays all the limits that you added in a **DiAGrammatic** view (as known as DAG view). Whenever you add a new limit, the system adds a yellow rectangle to the DAG view. This rectangle represents the limit you added. As you continue to add limits, the system adds corresponding rectangles to the DAG view.

These rectangles in the DAG view give you a graphical view of all the limits that you added till now.

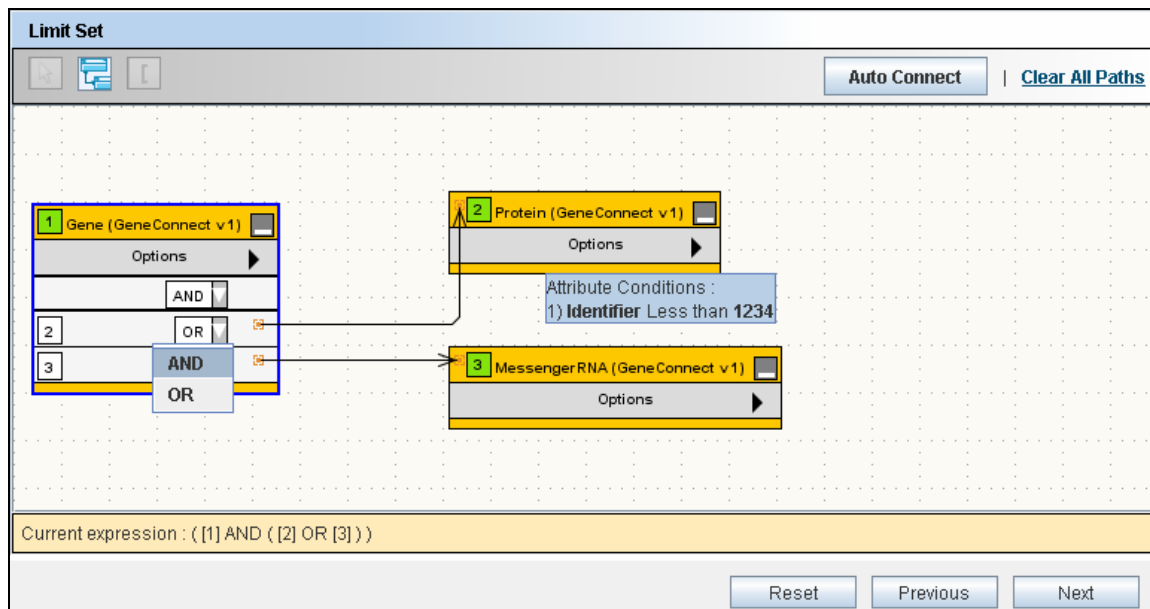
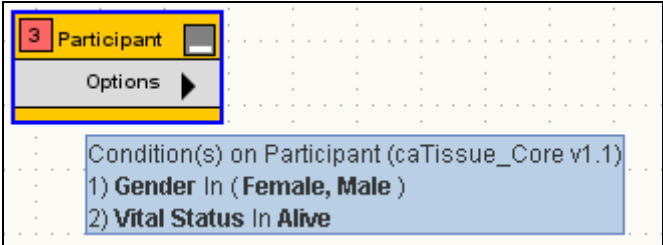
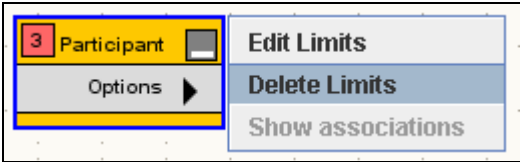


Figure 6 : Limit Set Pane

Operating on DAG

You can perform the following operations within the DAG view:

Table 5

| | |
|---------------------------|---|
| View details of the limit | <p>Position your mouse over a node to view the details of that limit.</p>  |
| Edit the limit | <p>Click Edit Limits from the Options menu of the node whose limits you wish to edit. Alternatively, double click the title of the node.</p>  <p>This opens the limit in the Define Limit Error! Reference source not found.pane in the edit mode. Edit the constraints in the Define Limit pane, and click the Edit Limit button.</p> |
| Delete the limit | <p>Select the nodes that you wish to delete and press DELETE.</p> <p>Alternatively, you can delete a single node by selecting Delete Limits from the Options menu of the node as shown above.</p> |

Connecting the Limits

You must connect all the limits before proceeding to the next step.

When you specify multiple limits, you need to connect these limits to form an expression. You can connect two limits using either an **AND** or an **OR** operator. The following is an example of how the DAG looks when the limits are connected.

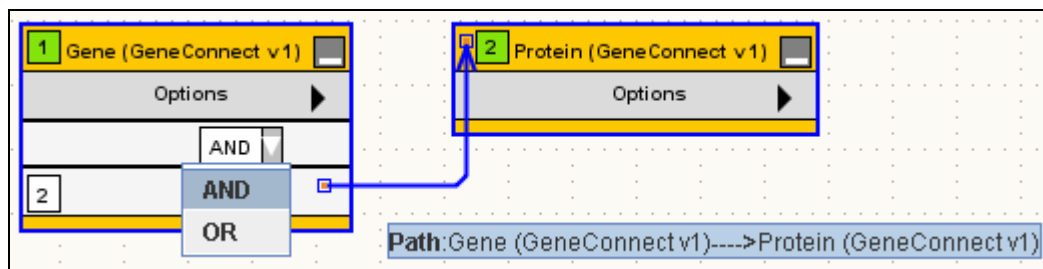


Figure 7: Connected Nodes

If there are only two limits in the query, you have to connect only those two rectangles

To connect the limits, you need to specify the following:

- The **path** to be taken in connecting the nodes
- The logical operator (**AND**, **OR**) in the dropdown box in a node

The rest of this section focuses on various ways in which you can specify the path(s).

What is a Path?

A path tells the system **how** two categories are to be connected. A path is a way to travel from one category to another. There could be multiple ways to connect two categories and the results of the query will differ based on the path you choose.

For example, consider the two categories *Specimen* and *Site* (from the caTissue Core application). A *Specimen* is associated with the *Site* category in two ways: 1) Site where the specimen was *collected* 2) Site where the specimen is *stored*.

That means there are two paths between *Specimen* and *Site*: 1) through Storage Container 2) through Specimen Collection Group for the collection site.

Query example 1: Show all the available specimens collected in Barnes Jewish Hospital. In this query, as you are specifically interested in a particular collection site, you should choose the path through Specimen Collection Group.

Query example 2: Show all the available specimens stored in Barnes Jewish Hospital. In this query, as you are specifically interested in a particular storage site, you must choose the path through Storage Container.

Therefore it is important to choose the right path based on what you are looking for.

Connecting Limits

Click the **Auto Connect** button. The system will automatically connect all the boxes based on the settings done by your administrator.

However, the administrator might not have configured the path between all the categories. In such cases, you can manually connect the limits as described below.

Example query: For all the male participants, show all the available specimens collected in Barnes Jewish Hospital.

1. After adding the three limits on *Participant*, *Specimen* and *Site* the DAG view will look like the figure below.

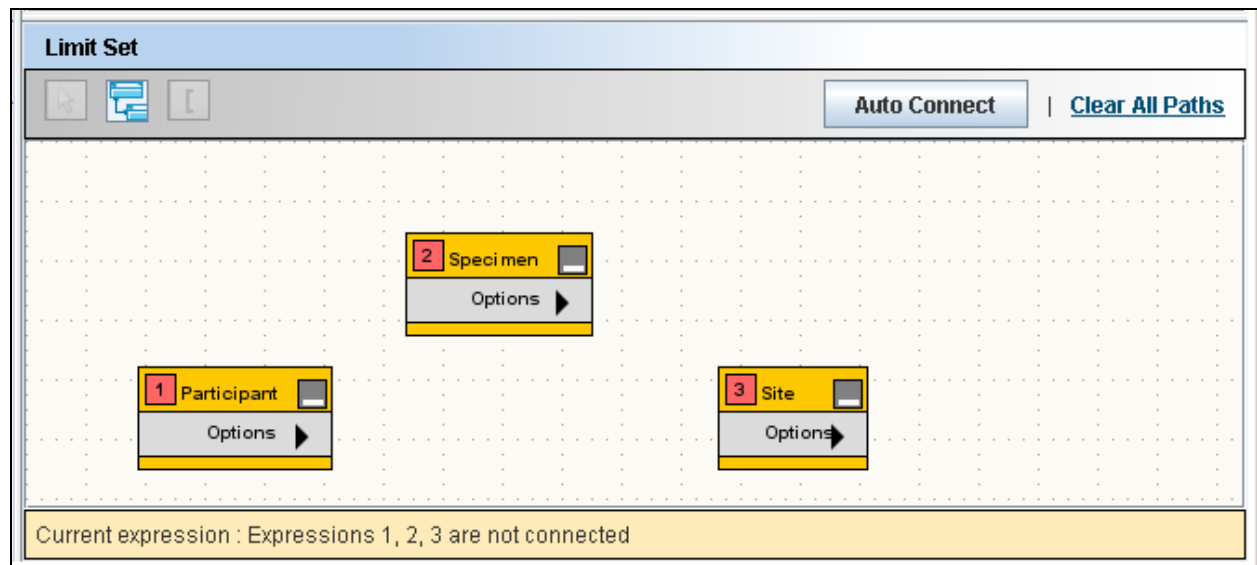



Figure 8: Limit Set

2. Click the *Participant* and the *Specimen* rectangles to select them.
3. Click the **Connect Limits**  icon to connect the two nodes. Now the DAG view will look like the figure below.

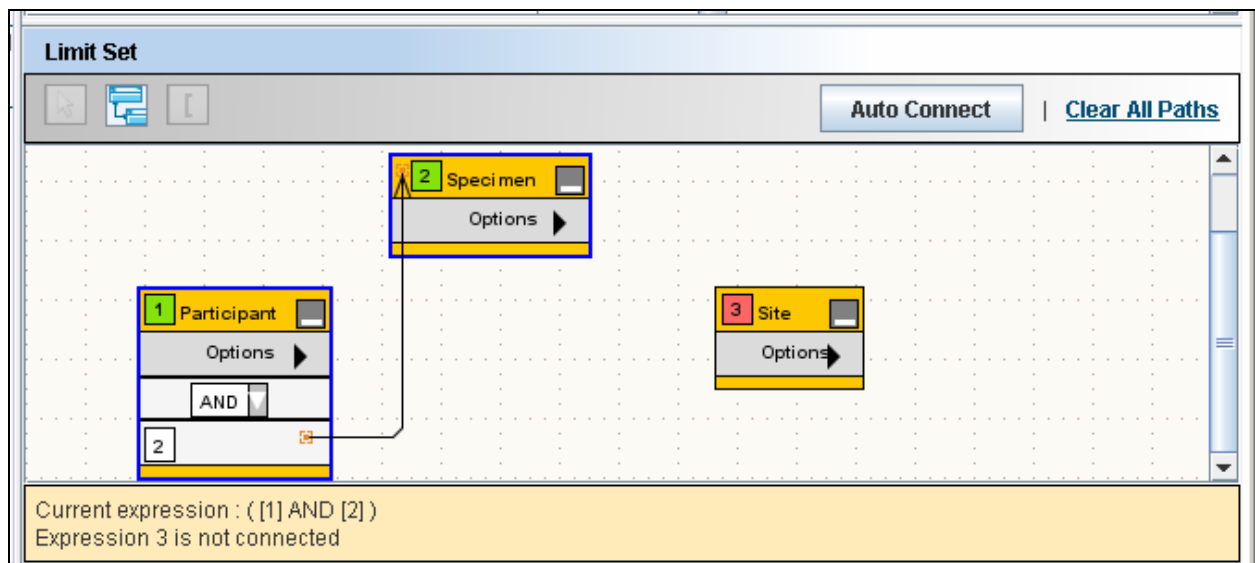



Figure 9: Limit Set with Two Connected Limits

Note: There is only one path between *Participant* and *Specimen*, therefore the system automatically connects it.

4. Select the *Specimen* and *Site* rectangles and click the **Connect Limits**  icon.
5. Since there are more than one paths between *Specimen* and *Site*, the system cannot connect the limits on its own. So it displays the **Ambiguity Resolver** window. The **Ambiguity Resolver** window looks like the figure below:

| Path Ambiguity Resolver | | |
|-----------------------------|--|----------------|
| General Path Curated Path | | |
| Select | Paths | PathPopularity |
| <input type="checkbox"/> | Specimen >> (specimenCollectionGroup) >> SpecimenCollectionGroup >> (site) >> Site | 50 % |
| <input type="checkbox"/> | Specimen >> (storageContainer) >> StorageContainer >> (site) >> Site | 50 % |

Figure 10: Path Ambiguity Resolver

The **Ambiguity Resolver** now shows two paths between *Specimen* and *Site*: 1) through Specimen Collection Group for the collection site 2) through Storage Container.

6. Select the first path to find out the collection site.
7. Once the path(s) are determined, the system creates a connection between the two nodes (see figure below). You can position the mouse over the connection to view the path(s) that you chose.

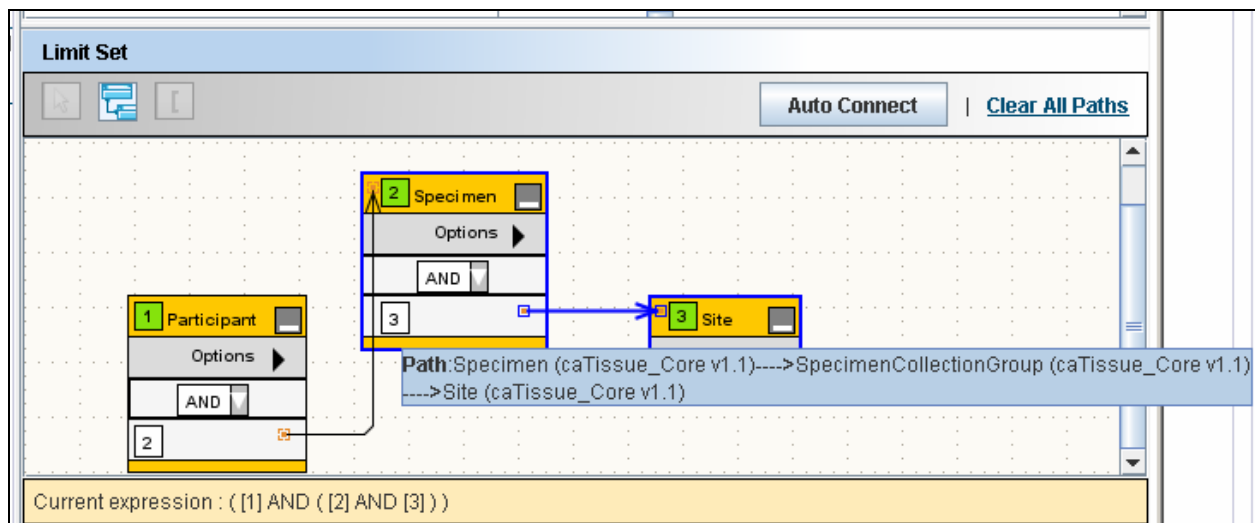


Figure 11: Limit Set with Three Connected Limits

8. Specify the operator (**AND** or **OR**) using the dropdown box in the node.
9. Click the **Next** button to go to the next step.

Deleting a Connection

You can delete a connection by clicking the connection and pressing the DELETE key.

The Information Pane

The **Information** pane is located at the bottom of the **Limit Set** pane.

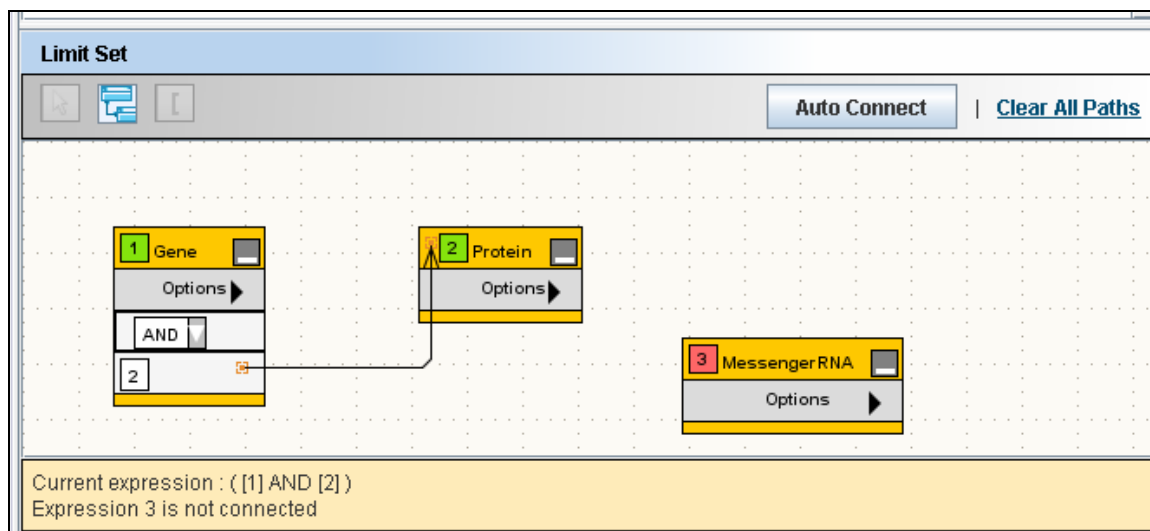


Figure 12: The Information Pane

The **Information** pane provides the following information:

- The expression formed as a result of the way you have connected the limits
- Limits that not yet connected to any other limit
- To execute the query, there should not be any unconnected rectangle.

Reset, Previous and Next Buttons

These buttons are located at the right hand bottom of the screen.

- The **Reset** button clears all the limits that you have specified.
- The **Previous** button takes you to **Choose Search Category** step of the wizard.
- The **Next** button takes you to the **Choose Output Category** step of the wizard. To proceed to the next step, the query that you have formed must be valid. If the query is invalid, an appropriate error message is displayed.

Define Search Results View

Specify the limits and then you need to choose the category you want to fetch the data for. For example, show all the participants who had biospecimens collected from their breast. In this case, *Participant* is the category that you want to view, based on a limit on the specimens.

The **Define Search Results View** tab enables you to select the output category for the query.

Figure 13: Define Search Result View.

All the categories on which you have specified limits are shown in the **Select Default View** dropdown box at the bottom of the screen. From the **Select Default View** dropdown box, select the category whose data you wish to obtain.

| | |
|--|--|
| | <p>Ignore the other contents of this screen such as the check boxes and hyperlinked categories. Those are intended to allow you to fetch data for multiple related categories in one go. But this feature is not supported in this release of caB2B.</p> <p>See the sections on In the image above, you can view the following sections:</p> <ul style="list-style-type: none"> Value of all the data elements in the record |
|--|--|

The Next and Previous Buttons

- To edit the limits, click the **Previous** button to go to the **Define Limit** step.
- Click the **Next** button to execute the query that you have created.

| | |
|----------------|--|
| CAUTION | <p>On clicking the Next button, the system will execute the query and fetch the data from appropriate data services over the caGrid.</p> <p>This step could be time-consuming based on the network speed and the complexity of the query.</p> |
|----------------|--|

View Search Results

The **View Search Results** tab displays the results obtained by executing the query you had defined in the previous steps. The concept of viewing results in caB2B is similar to viewing results after a Google search. The similarities are as follows:

- caB2B returns a set of records that match your limits.
- Some details of the record is displayed along with the record.
- You can click on one of the records to view more details.
- There you can view additional associated information for the record you are viewing. For example, from Participants you can see the Collection Protocol in which they registered. This is similar to Google where, after you click one of the hyperlinks, you can traverse into the other web pages by clicking the hyperlinks within that page.
- As in Google, the results are divided into multiple pages for easy navigation.
- If you find any data of your interest, you can add it to a shopping cart known as **Data List**.

The following image displays a typical results view. The **View Search Results** tab displays the records that match the limits and also the data that you have already added to the data list.

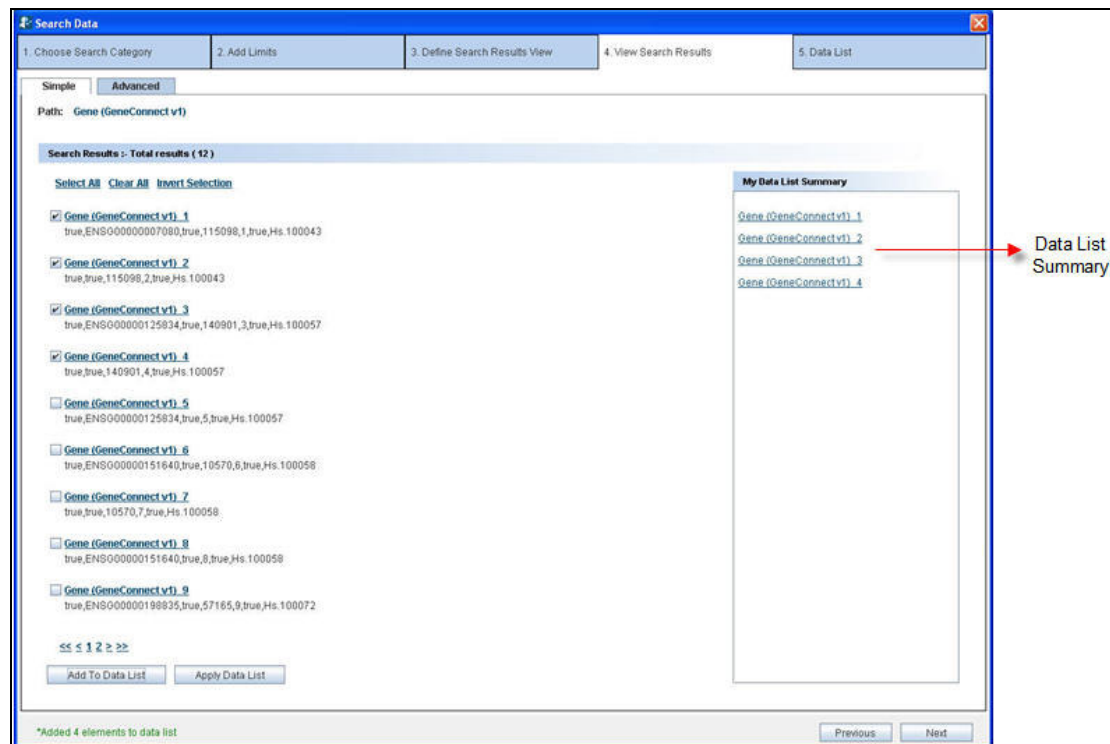


Figure 14: View Search Results

Record Details

Click on the hyperlink of a record on the **View Search Results** to take you to the **Record Details** page.

Here you can view the values of all the attributes of the category.

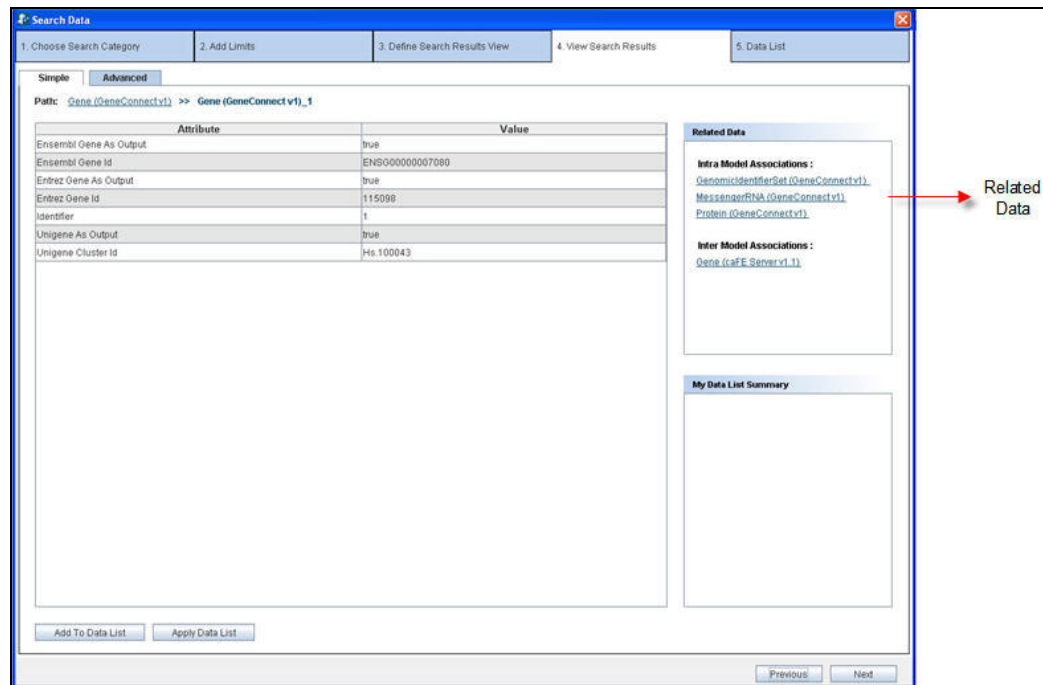


Figure 15: Results' Details

In the image above, you can view the following sections:

- Value of all the data elements in the record
- Links to related data
- Summary of the records that you have already added to the data list

Related Data

The **Related Data** box on the right displays the categories that are related to the current category. Click any of them to display the records for that category. For example, if you are currently viewing records of Gene(GeneConnect), the following categories are shown in the related data box:

- GenomicIdentifierSet (GeneConnect)
- Protein (GeneConnect)
- MessengerRNA (GeneConnect)
- Gene (caFE Server)

If you now click *Protein*, you would see the Proteins related to the Gene you are currently viewing.

Breadcrumbs

Definition of Breadcrumbs – Wikipedia

Breadcrumbs or **breadcrumb trails** are a navigation technique used in user interfaces. Its purpose is to give users a way to keep track of their location within programs or documents.

As you traverse the results view, breadcrumbs are displayed at the top of the page. For example, if you are viewing a Protein record related to a Gene, you would see the path as

Gene >> Gene _1 >> Protein >> Protein_5

This means that you originally queried for Gene. Then you fetched related proteins for Gene_1. From the proteins obtained, you are currently viewing details of Protein_5.


Click any of the links in the breadcrumbs to navigate back to the corresponding page. For example, you can click *Gene* in the above path to go back to the page where you were shown the list of Genes.

Adding data to the Data List

A data list is like a shopping cart. As you traverse through the results and find some interesting data you can add it to the shopping cart.

You can add data to the data list in two ways:

- Click the **Add to Data** List button
- Click the **Apply to Data** List button

| | |
|---|---|
|  | As you add the data to the data list, they are reflected in the My Data List Summary box on the right hand side. |
|---|---|

Add to Data List

Select the record that you want to add to the data list and click the **Add to Data List** button. You can either add one record at a time or add multiple records by selecting the checkboxes placed along with the record summary.

Using Apply Data List

Why use Apply Data List

As explained in the **In the** image above, you can view the following sections:

- Value of all the data elements in the record
- Links to related data
- Summary of the records that you have already added to the data list

Related Data section, you can obtain data for categories related to a particular category.

Suppose you do the following:

1. Fetch the data for the *Gene* category.
2. Fetch the related data from the category *Chromosome* for a record of *Gene*, for example, *Gene_1*.

3. Add the *Chromosome* record to the data list.
4. The data list now looks like the figure below.

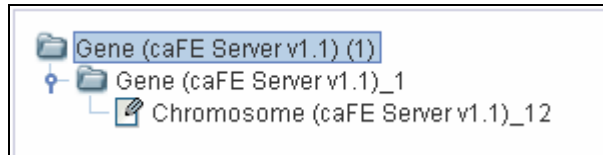


Figure 16: Data List

Now if you wish to obtain the *Chromosome* data for some more *Gene* records you have to fetch the related *Chromosome* for each *Gene* record one at a time. The **Apply Data List** functionality allows you to fetch the *Chromosome* for several *Gene* records in one go. To do this (continuing the previous steps)

To fetch the *Chromosome* for several *Gene* records in one go, perform the steps 1,2, and 3. Now perform the following steps:

5. From the **View Results** screen for the *Gene* records, (see Figure 14: View Search ResultsError! Reference source not found.) select *Gene* records for which you want the related *Chromosome* records.
6. Click the **Apply Data List** button.
7. A message is displayed in the status bar when the operation is complete. Click **Next** to see the data list. The data list now looks like the figure below.

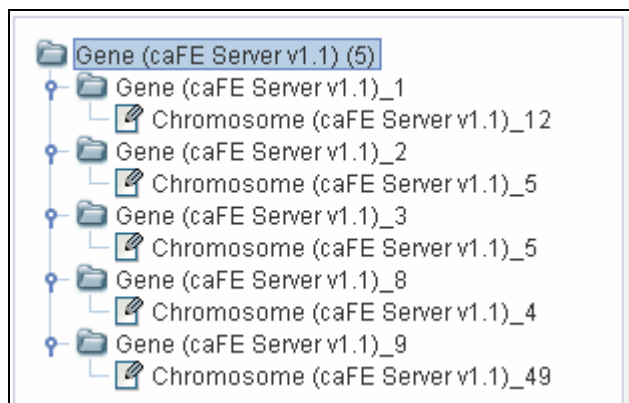


Figure 17: Result of Apply Data List to genes with id 2, 3, 8, 9

Thus, once you obtain the related data for a particular record, **Apply Data List** allows you to fetch similar related data for several other records easily.

How the Apply Data List functionality works

To understand this functionality better, consider a slightly more complicated data list below:

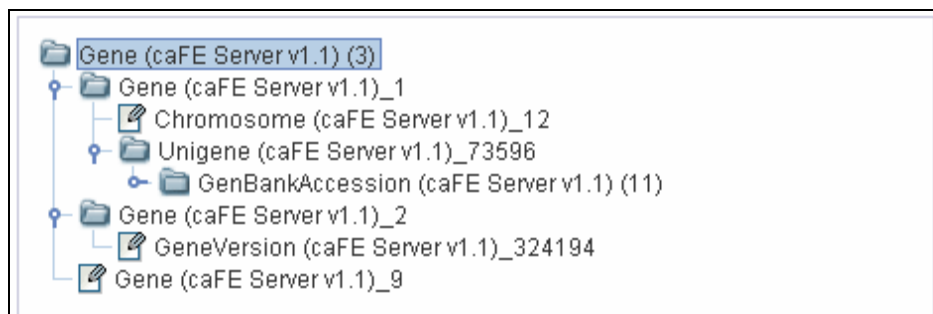


Figure 18: Example 3- Data List

Example 3

The following records have been added to the data list

- For *Gene_1*, related *Chromosome*, *Unigene*, and *GenBankAccession* records are present.
- For *Gene_2*, related *GeneVersion* is present.
- For *Gene_9*, there is no other related data.

Now, **Apply Data List** to the genes *Gene_1*, *Gene_2* and *Gene_9*. The resulting data list is shown below:

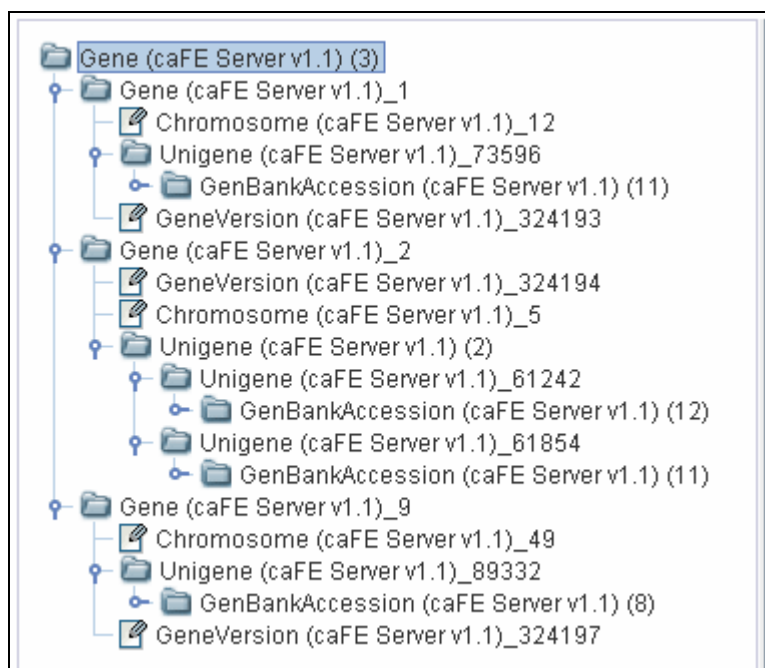


Figure 19: Example 3- Apply Data List

As seen from above figure, following new related data has been fetched:

- For *Gene_1*, the system fetches *GeneVersion* (as *Gene_2* had *GeneVersion*).
- For *Gene_2*, the system fetches *Chromosome*, *Unigene*, and *GenBankAccession*, as these were the related data for *Gene_1*.
- For *Gene_9*, the system fetches all of *GeneVersion*, *Chromosome*, *Unigene*, and *GenBankAccession*.

As a result, each gene contains the related data for *GeneVersion*, *Chromosome*, *Unigene*, and *GenBankAccession*.

Thus, the **Apply Data List** operation results in a homogenous structure in the data list. The system fetches related data for all the other records like the related data present for one record.

The Next and Previous Buttons

- Click **Next** to view the contents of the data list.
- Click Previous to go back to Choose Search Category tab.

Viewing and Saving Data List

After adding the records of your interest to the data list, you can get the consolidated view of the data list. You can also export the data list into a CSV file or create an Experiment using the contents of the data list.

The following figure shows the data list tab:

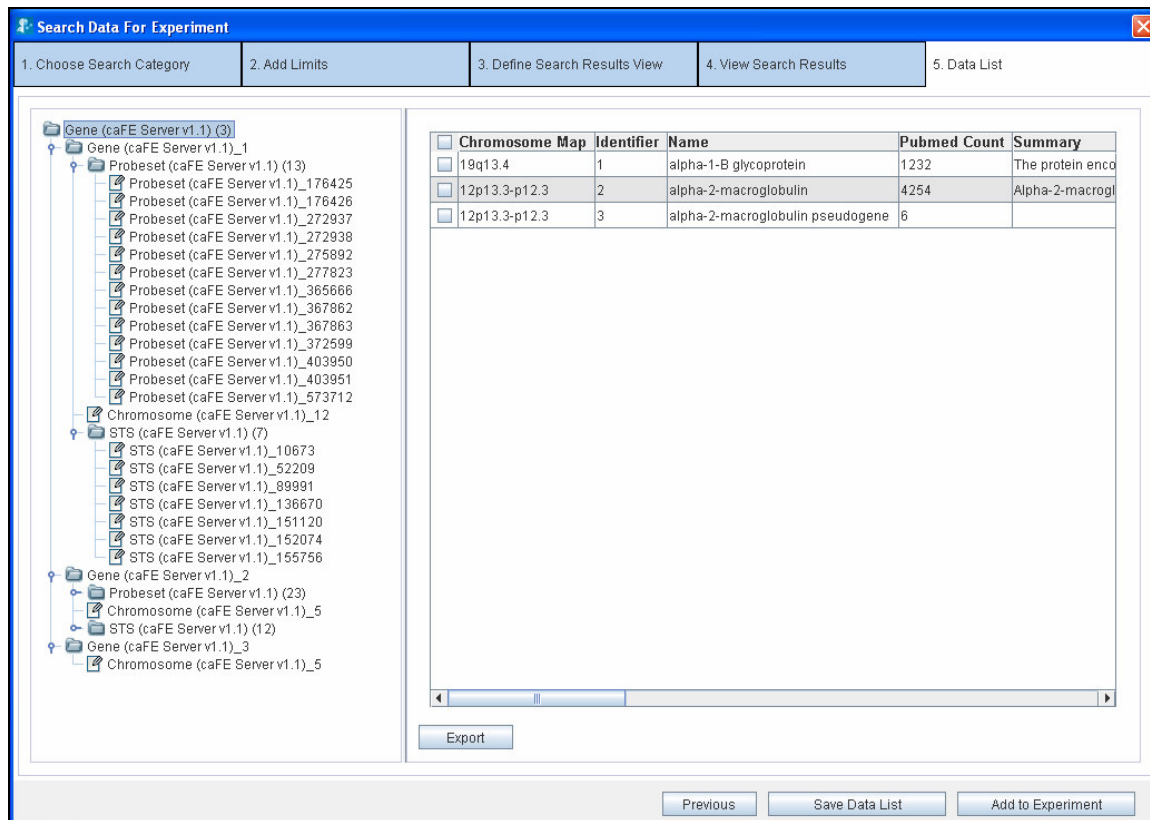


Figure 20: Data List

The left-hand side pane shows all the records that you have added to the data list in a tree format. Each node in the tree represents either a category or a record of a category.

- For a node representing a category:
 - The number of records for that category is shown in parentheses. For example *Gene (3)* means that you have added 3 records for the *Gene* category.
 - Clicking on the node updates the right hand side pane to show a spread sheet containing the records of that category (see figure above).
- For a node representing a record
 - The unique identifier for that record is displayed. For example, *Gene_1* represents the record with identifier 1.
 - Click the node updates on the right hand side pane to show the details of that record.

Operations on the Data List

You can do the following operations on the data list:

- Save the data list – Click the **Save Data List** button.
- Export the data list into a CSV file – Click the **Export** button.
- Add the data list to an Experiment – Click the **Add To Experiment** button.

To save the data list:

1. Click the **Save Data List** button.
2. Type the title and description for the data list.
3. Click the **Save** button. The system displays a message in the status bar indicating that the data list is saved.

Chapter 3 Experiments

In this Chapter

After you search and add some data of interest to the data list, you can create an experiment using that data. This chapter explains the various operations that you can perform on the data sets that you collected.

The following are the sections in this chapter:

- What are Experiments and Projects?
- Creating an Experiment
- Viewing the Existing Experiments

Working with an Experiment

- Filtering Data
- Analyzing Data

What are Experiments and Projects?

An experiment is a collection of data. After you create an experiment, you can perform the following:

- Perform various analyses on the data using the grid-enabled analytical services
- Visualize the analysis results using various viewers such as bar charts, scattered plots and so forth
- Create new data sets by applying filters on the existing datasets

A project is a collection of experiments. A project can contain many Experiments or other sub-projects. The grouping of experiments into projects helps you manage the multiple experiments easily.

Example of an experiment

Aim: Find out the micro array data for a nonmalignant brain tissue. Perform **Comparative Marker Selection (CMS)** on this data to determine the expression values that are most closely correlated with a class template. View the correlation among the expression values as a scattered plot.

You need to perform the following steps using caB2B:

| Step no. | Step | Related Section |
|----------|--|--------------------------------|
| 1 | Search for micro array data | Search data – The Query Wizard |
| 2 | Create experiment with this data | Creating an Experiment |
| 3 | Open the experiment you have created | Opening an Experiment |
| 4 | Filter the data further if desired | Filtering Data |
| 5 | Save the filtered data | How to save filtered data? |
| 6 | Perform CMS analysis on selected data | Analyzing Data |
| 7 | Plot analysis result as a scattered plot | Visualizing Data |

Creating an Experiment

To create an experiment, you need to have a saved data list with the required data. See Viewing and Saving Data List for details on how to do this.

After you have saved the data list, do the following to create an experiment:

1. Click the **Add to Experiment** button (Figure 20: Data List). This opens the **Create New Experiment** window.

Create New Experiment

* Experiment Name : FE gene Trials

* Project :

Add New

- My Projects
 - Gene Analysis
 - MicroArrays
 - Cell Free Nucleic Acid Detection
 - Texture Analysis
 - Gene Analysis
 - Participants analysis
 - Participants analysis

Description : This is the experiment on genes and chromosomes from FE application

Cancel Save

Figure 21: Create New Experiment

2. Give an appropriate name to the experiment.
3. Select the project in which you wish to create this experiment (see note below).
4. Provide a description of the experiment.
5. Click **Save**.

You can now close the **Search Data** wizard to return to the **Home** page and open the experiment.

| | |
|--|--|
| | <ul style="list-style-type: none"> • To create a new project, perform the following: <ol style="list-style-type: none"> 1. Select the parent project under which you wish to create the new project 2. Click the Add New button. This adds a new sub project under the selected project with the default name— <i>New Project</i>. • To rename the new project <ol style="list-style-type: none"> 1. Select the project. 2. Press F2. 3. Type the desired name. 4. Press ENTER. |
|--|--|

Viewing the Existing Experiments

On the caB2B Home page, click the Experiment tab. This opens the My Experiments page.

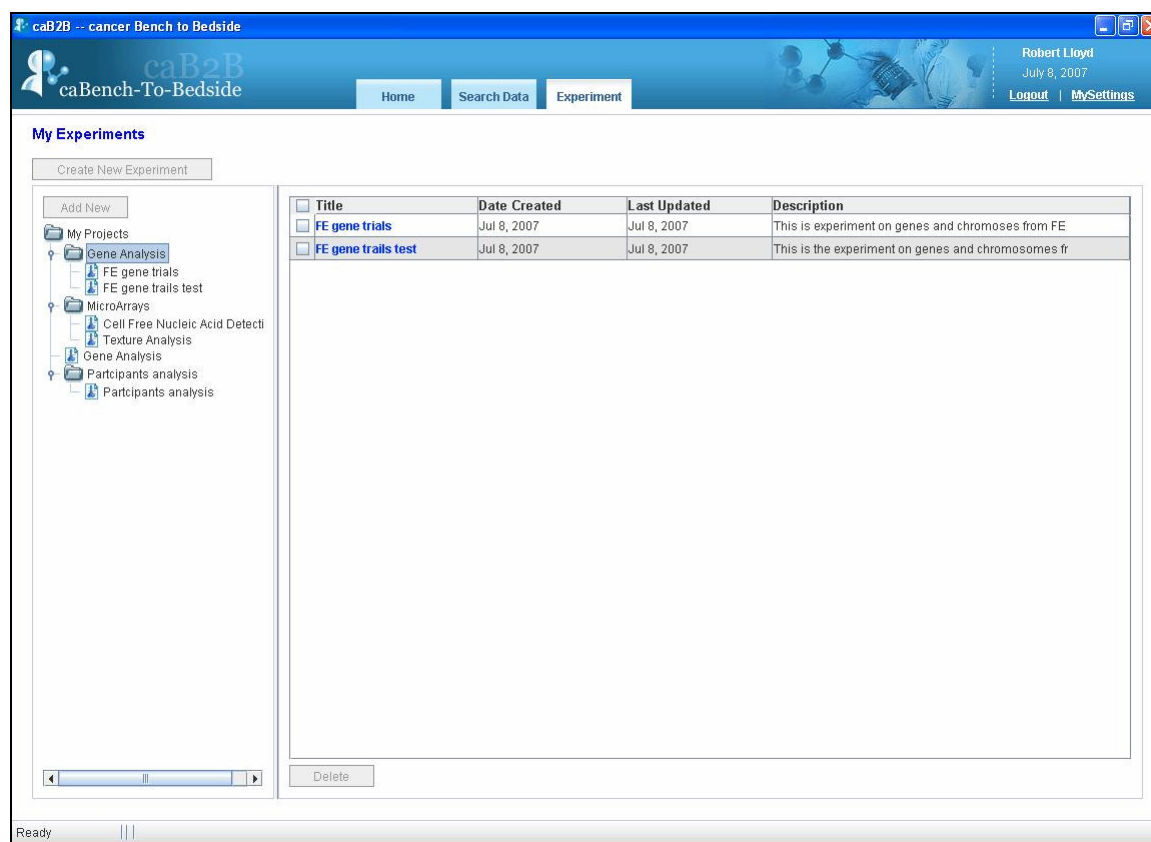


Figure 21: My Experiments

You can see all the projects and experiments that you created in the left hand side pane.

When you select a project in the left hand side pane, the system displays a list of the experiments in the project in right hand side pane.

Opening an Experiment

To open an experiment, click the title on that experiment on the right hand side pane of the My Experiments page (see Figure 21: My Experiments).

When you open an experiment, you see the Experiment Details page.

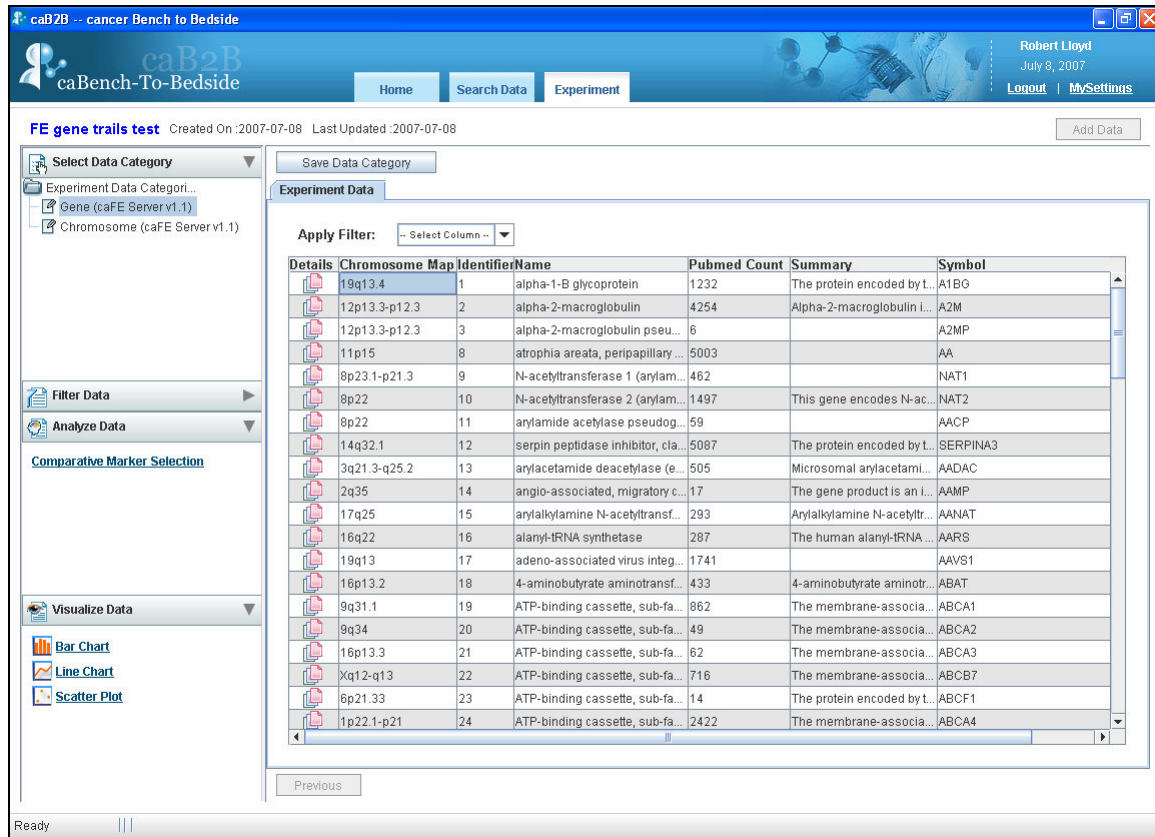


Figure 22: Experiment Details

Working with an Experiment





The **Experiment Details** page consists of the following sections:

- The **Operations** pane on the left hand side
- The **Visualization** pane on the right hand side

The Operations pane

The **Operations** pane contains the following collapsible boxes.

Table 1: Operations on the Experiment Page

| UI Elements | Description |
|---|---|
|  Select Data Category ▼ | Displays a list of all the data categories present in the experiment. See <i>How to save filtered data?</i> |
|  Filter Data ▼ | Shows all the filters that are currently applied to the data of the selected data-category. See <i>Filtering Data</i> for details. |
|  Analyze Data ▼ | Shows the links to call analytical services that are applicable to the selected data-category. See <i>Analyzing Data</i> for details. |
|  Visualize Data ▼ | Shows the visualization options available for the selected data-category. See <i>Visualizing Data</i> for details. |

Visualization pane

This is the right hand side pane on the **Experiment Details** page (see Figure 22: Experiment Details). This pane can contain several pages at a time (see figure below). You can switch between the pages by clicking on the **tabs** at the top of this pane.

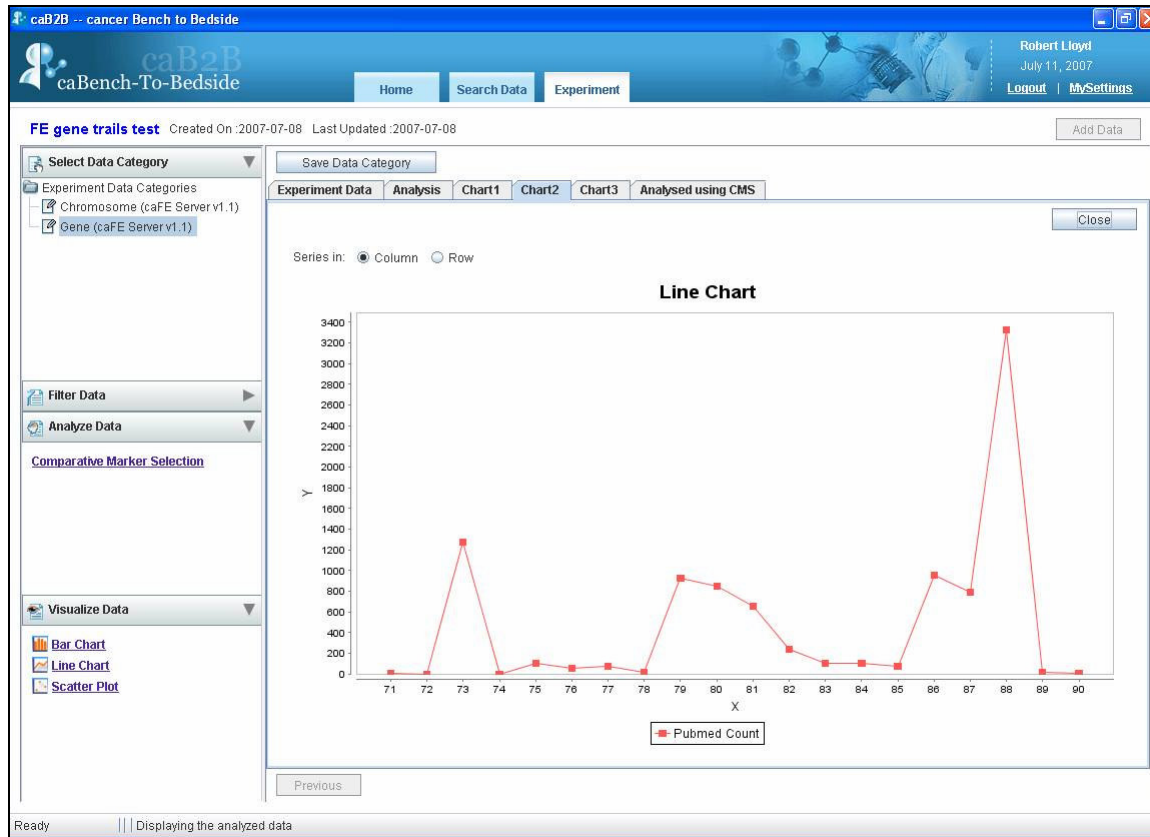


Figure 23: Visualization Pane

The visualization pane is updated based on the operations you do in the operations pane. For example,

| Operations pane | What happens in Visualization pane |
|--|---|
| Click a data-category in the Select Data Category box | The Experiment Data tab is updated to show the records of that data-category |
| Choose a visualization from the Visualize Data pane | A new tab is created to display the desired visualization |

Filtering Data

The purpose of creating experiments is to perform analysis of the data. To aid the analysis, caB2B provides features to filter the data so that you can zero in on the data of your interest.

After you filter the data, you can save the filtered data as a **data-category**. For example, if you have created an experiment with all female Participants, you can now create different data categories based on age less than 50, age between 50 and 65 and age greater than 65.

How to filter data?

1. From the **Apply Filter** dropdown list, (located at the top of the **Experiment Data** tab), select the

column of whose values you wish to filter the data.

2. System displays appropriate filter controls.
3. Specify the values that the filtered data must satisfy.
4. The data in the **Experiment Data** tab is updated to show the filtered data.
5. The filter that you have applied is also shown in the **Filter data** box in the Operations pane on the left hand side.

You can now save the filtered data as a new data-category and use it to perform analyses or visualization.



You can filter on multiple columns by successively applying filters on those columns. For example, to obtain *female participants between age 30 and 40*

1. Apply the filter “Gender = female”.
2. Then apply the filter “age between 30 and 40”.

For details of the various filter controls and how to use them, see Types of Filters.

How to edit a filter?

1. From the **Apply Filter** dropdown, choose the column whose filter you wish to edit.
2. An appropriate filter control is shown with the values that you previously entered.
3. Specify desired values of the filter (see Types of Filters).
4. The data in the **Experiment Data** tab is updated to show data based on the new filter.
5. The filter criteria shown in the **Filter data** box in the operations pane is also updated.

How to remove a filter?

1. From the **Apply Filter** dropdown, choose the column whose filter you wish to remove.
2. An appropriate filter control is shown with the values that you previously entered.
3. Remove the values that you had specified (see Types of Filters).
4. The data in the **Experiment Data** tab is updated to show data on removing the filter.
5. The filter is also removed from the **Filter data** box in the operations pane.

How to save filtered data?

Once you apply filters to a data set, you can save the filtered data as a **data-category**. You can create as many data-categories as you want. The data-categories can be further used for data analysis.

Once you have obtained filtered data, you might wish to save it as a new data-category.

To do this,

1. Click the **Save Data Category** button (located at the top of the Visualization pane).

2. The **Save as Data category** pane is displayed.
3. In the **Title** field, type the title of the new data-category.
4. Click **Save**.

The new data-category will appear under **Custom Data Categories** section in **Select Data Category** box of the operations pane.

Analyzing Data

After you searched the data you are interested in and created experiments from it, you can analyze that data. To do this you can use various analytical services available. Typically they accept data and perform some kind of analysis on it and return the result. You can also pass some configuration parameters along with the data to the analytical services.

When you are viewing the data, the services that can operate on that data are shown in the **Analyze Data** box. In the current version one analytical service **Comparative Marker Selection**, from gene pattern is available.

How to analyze data?

1. Select the data category from **Select Data Category** box, for example *Bioassay*.
2. In the **Analyze Data** box, click on the link to call the service, for example **Comparative Marker Selection** link.
3. It will open a popup that accepts configuration parameters for the service. Figure 24: Configuration parameters for CMS shows such parameters for **CMS**
4. Enter the **Title** for this analysis and values for the various parameters
5. Click **Finish** button to submit this analysis.

Comparative Marker Selection Parameter Set ✕

Analysis Title :

Balanced : ☒ True ☐ False

Complete : ☐ True ☒ False

Min Std :

Number Of Permutations :

Phenotype Test : ▼

Random Seed :

Smooth Pvalues : ☐ True ☒ False

Test Direction : ▼

Test Statistic : ▼

Figure 24: Configuration parameters for CMS

You can see the new **Analysis** tab added in the right hand side panel. It contains following fields:

Experiment Data

Analysis

Analysis performed for 'FE gene trails test'

| Data Category | Analysis Title | Date | Status |
|-------------------------|---------------------------------|-------------|-----------|
| Gene (caFE Server v1.1) | CMS for Hu95Av2 | 12-Jul-2007 | Completed |
| | | | |

Figure 25: List of analysis performed for the experiment

| Field | Description |
|----------------|--|
| Data Category | The data category on which you performed the analysis. |
| Analysis Title | The title you provided for the analysis. It is a link to the result of the analysis. |
| Date | The date on which you performed the analysis. |
| Status | Indicated whether analysis is completed or pending . |

You can view the result on **completed** analysis by clicking the link in Analysis **Title** field.




















| Experiment Data Analysis CMS for Hu95Av2 | | | | | |
|---|----------------|------------|--|--------------|--|
| Close | | | | | |
| Apply Filter: -- Select Column -- | | | | | |
| Details | Chromosome Map | Identifier | Name | Pubmed Count | Summary |
|  | 19q13.4 | 1 | alpha-1-B glycoprotein | 1232 | The protein encoded by this gene is a plas |
|  | 12p13.3-p12.3 | 2 | alpha-2-macroglobulin | 4254 | Alpha-2-macroglobulin is a protease inhibi |
|  | 12p13.3-p12.3 | 3 | alpha-2-macroglobulin pseudogene | 6 | |
|  | 11p15 | 8 | atrophia areata, peripapillary chorioretinal ... | 5003 | |
|  | 8p23.1-p21.3 | 9 | N-acetyltransferase 1 (arylamine N-acetyltr... | 462 | |
|  | 8p22 | 10 | N-acetyltransferase 2 (arylamine N-acetyltr... | 1497 | This gene encodes N-acetyltransferase 2 (|
|  | 8p22 | 11 | arylamine acetylase pseudogene | 59 | |
|  | 14q32.1 | 12 | serpin peptidase inhibitor, clade A (alpha-... | 5087 | The protein encoded by this gene is a plas |
|  | 3q21.3-q25.2 | 13 | arylacetamide deacetylase (esterase) | 505 | Microsomal arylacetamide deacetylase con |
|  | 2q35 | 14 | angio-associated, migratory cell protein | 17 | The gene product is an immunoglobulin-ty |
|  | 17q25 | 15 | arylalkylamine N-acetyltransferase | 293 | Arylalkylamine N-acetyltransferase belongs |
|  | 16q22 | 16 | alanyl-tRNA synthetase | 287 | The human alanyl-tRNA synthetase (AARS) |
|  | 19q13 | 17 | adeno-associated virus integration site 1 | 1741 | |
|  | 16p13.2 | 18 | 4-aminobutyrate aminotransferase | 433 | 4-aminobutyrate aminotransferase (ABAT) i |
|  | 9q31.1 | 19 | ATP-binding cassette, sub-family A (ABC1)... | 862 | The membrane-associated protein encode |
|  | 9q34 | 20 | ATP-binding cassette, sub-family A (ABC1)... | 49 | The membrane-associated protein encode |
|  | 16p13.3 | 21 | ATP-binding cassette, sub-family A (ABC1)... | 62 | The membrane-associated protein encode |
|  | Xq12-q13 | 22 | ATP-binding cassette, sub-family B (MDR/... | 716 | The membrane-associated protein encode |
|  | 6p21.33 | 23 | ATP-binding cassette, sub-family F (GCN2... | 14 | The protein encoded by this gene is a merr |

Figure 26: Result of CMS Analysis

Visualizing Data


You can visualize the data in the experiment using various graphical components like charts and scatter plot. These options provides graphical summary of the data and helps in comparing it, finding co-relation and variation in it.

Some Terms that are used:

Data series: Related data points that are plotted in a chart and originate from data rows or data columns. Each data series in a chart has a unique color or pattern. This is represented by a *Legend*. You can plot one or more data series in a chart.

Legend: A small color box that identifies each series in the chart.

The current version of caB2B supports following visualization options in **Visualize Data** box:

| Visualize option | Description |
|--|--|
|  Bar Chart | This is a chart with rectangular bars of lengths proportional to the data value it represents. |




| | |
|--|---|
| | |
|  Line Chart | This is a chart with a line joining the points that represents data value |
|  Scatter Plot | A scattered plot will show up a linear or non-linear relationship between the variables (column selected). It also helps in finding odd-mans in the data. |

Table 2: Types of charts supported

How to plot data series:

1. Select the data that you want to plot from **Experiment tab**.
2. Click on the link for type of the chart you wanted, for example  [Bar Chart](#)
3. It will open the chart in a new tab in right hand side panel.
4. By default data series are from columns on the **Experiment Data** tab.

Following screen shots represents the various charts for **pubmed count** for few Genes from FE server).

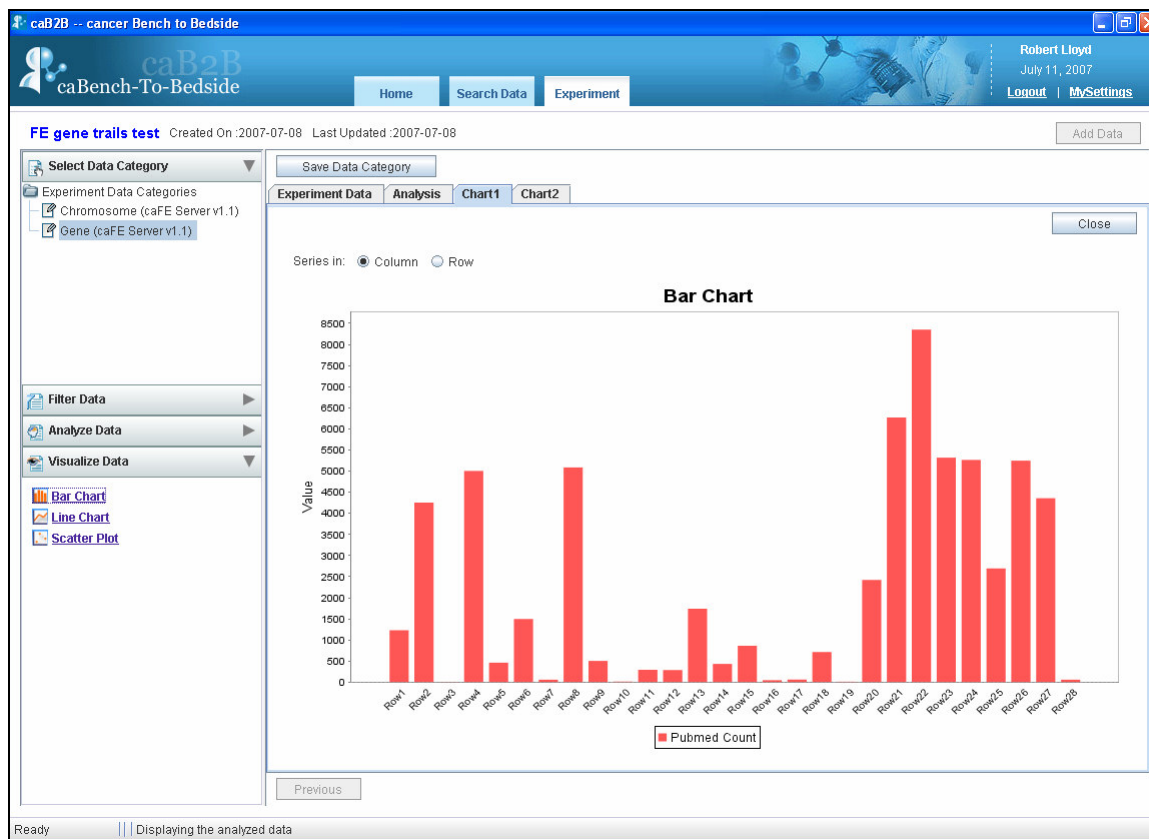


Figure 27: Bar chart example

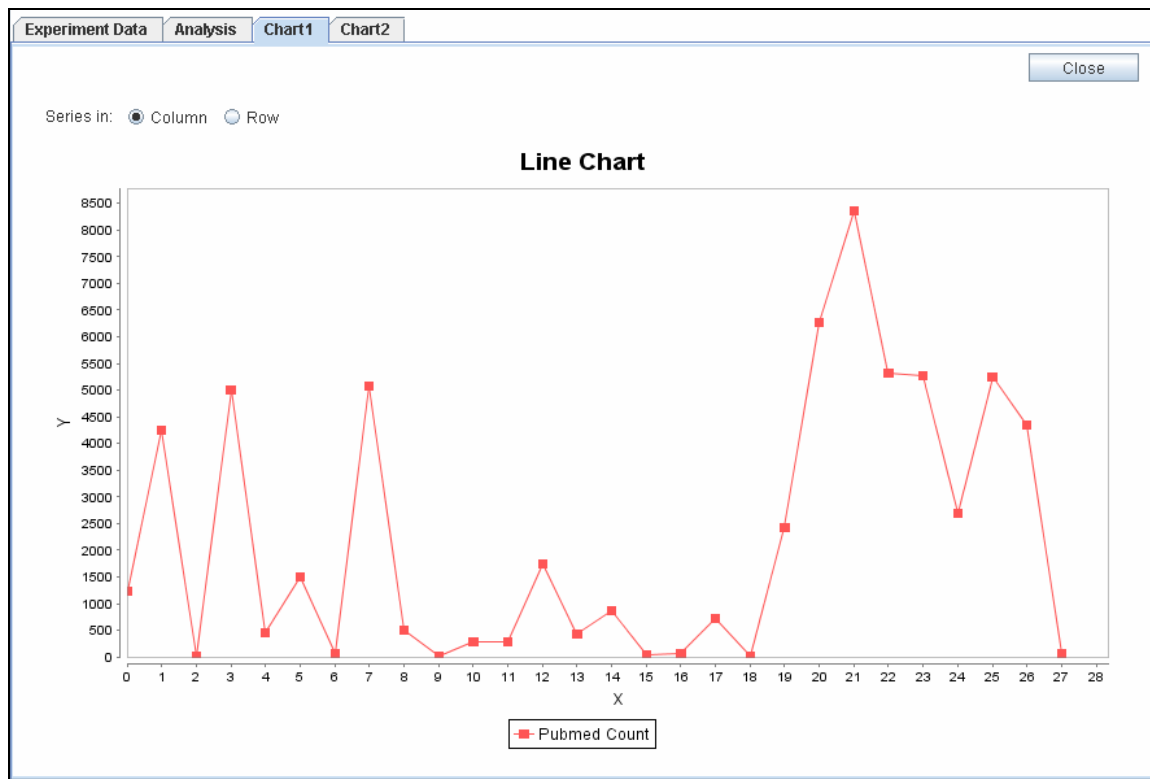


Figure 28: Line chart example

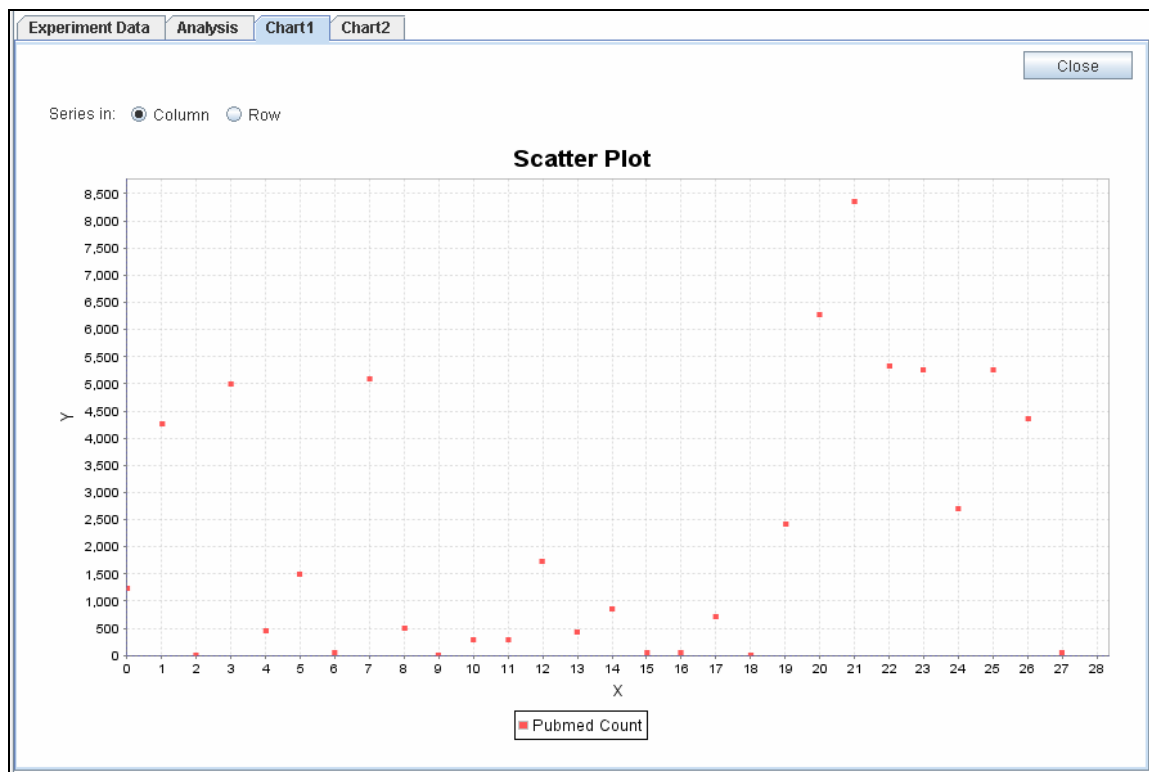


Figure 29: Scatter Plot example

How to change data series from rows to column and vice versa?

You can configure whether data series are from rows or columns by selecting appropriate **Series In** options.

- To have the data columns to represent the data series, select ☒ Column option.
- To have the data rows to represent the data series, select ☐ Row option.

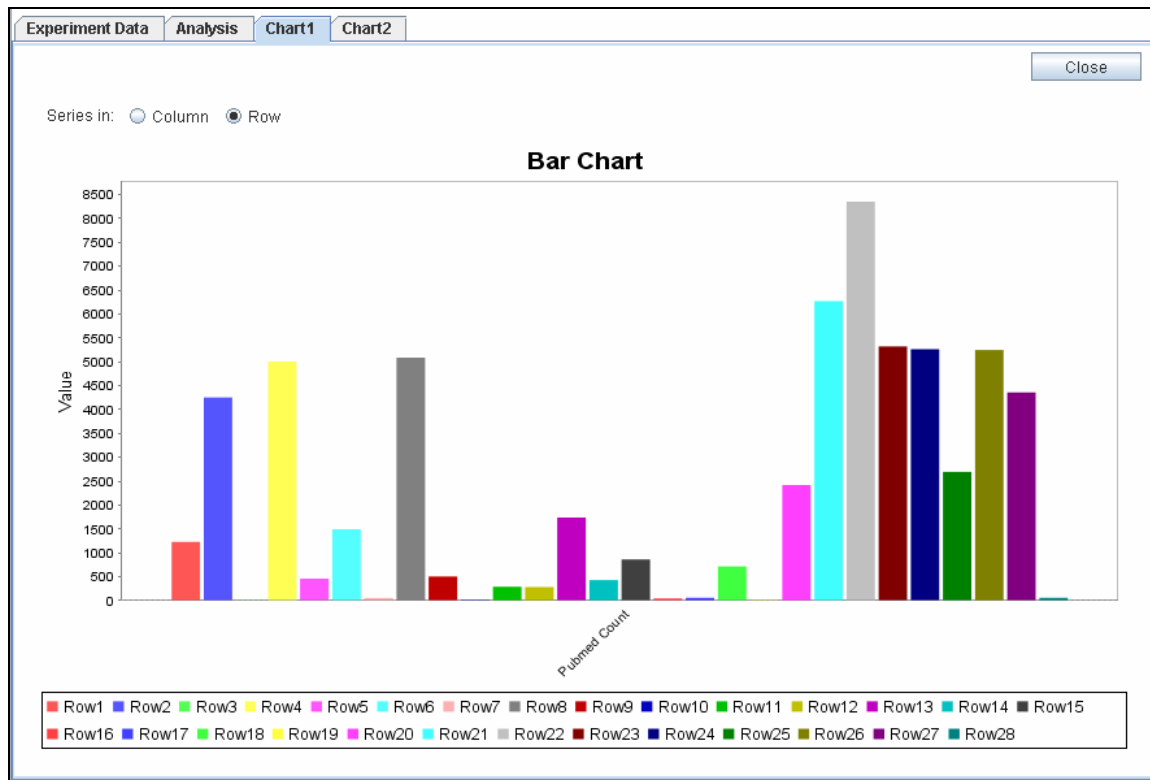


Figure 30: Bar chart with data series from rows



You can change the type of the chart you are currently viewing by selecting required chart link from **Visualize Data** box.

You can create multiple charts for the data.

The name of the tab added for each chart is “*chart*” followed by a unique number for that chart. For example *chart1*, *chart2* and so on.

Chapter 4 Appendix 1

Examples of category search

This appendix provides examples of the advanced search when you search for categories.

Text-based Search - Example 1

To search categories having *gene* in their name, you need to do following:

1. In the **Search** box, type *gene*.
2. Select the **Category** check box.
3. Click the **Text** option button.
4. Click the **Search** button.

The results of this search are as shown below. All the categories shown below have *gene* in their names.

(E.g. Participant, Gene, Experiment, Protein, Specimen)

Advanced Search

☒ Category ☐ Include Description
☐ Attribute ☐ Permissible Values
☒ Text ☐ Concept Code

Search Results :- Total results (16)

[Gene Annotation](#)
 Gene annotation includes information about a gene such as its symbol, gene name(s), c....

[Genomic Identifiers](#)
 Genomic identifiers from many publicly available databases (e.g. Entrez, UniGene, Gen....

[Literature-based Gene Association](#)
 Genes that are co-mentioned in the same PubMed abstracts have been curated. Each pai...

[Mi](#)
 Mi Genes that are co-mentioned in the same PubMed abstracts have been curated. ro...
 Each pair of genes as well as their correlation, calculated by squaring the
 Or number of abstracts where the two genes are co-mentioned divided by the product
 Ca of the number of abstracts mentioning each of the genes, is reported. f....

[EnsemblGene \(GeneConnect v1\)](#)
 This class representing the Ensembl Gene ID.

<< < 1 2 3 > >>

Figure 31: Search Result of Example 1

Text-based Search - Example 2

To search the categories having `Protein` in their name or description or to search the attributes having `Protein` in their name or description, you need to do following:

1. In the **Search** box, type **Protein**
2. Select the **Category** check box.
3. Select the **Attribute** check box.
4. Select the **Include Description** check box.
5. The **Text** option button is selected by default.
6. Click the **Search** button.

The result of this search is as shown in the screenshot below.

The screenshot displays a web-based search interface. At the top, there is a search input field containing the text 'protein' and a 'Search' button to its right. Below the input field, a hint text reads '(E.g. Participant, Gene, Experiment, Protein, Specimen)'. An 'Advanced Search' panel is open, showing several options: 'Category' and 'Include Description' are checked; 'Attribute' is checked and 'Permissible Values' is unchecked; 'Text' is selected with a radio button, while 'Concept Code' is unselected. Below this panel, a blue header bar indicates 'Search Results :- Total results (11)'. The results are listed as links with brief descriptions: 'Gene Annotation' (includes information about a gene), 'Genomic Identifiers' (from various databases), 'Microarray Annotation' (information about a probe), 'BioMaterial (caArray v1.1)' (abstract class for substances), 'BioSequence (caArray v1.1)' (representation of DNA, RNA, or protein), and 'GenBankAccession (caFE Server v1.1)' (GenBank Accession for a Gene). At the bottom, a pagination control shows '<< < 1 2 > >>' with the number '2' highlighted in red.

Figure 32: Search Result of Example 2

Concept Code Based Search - Example 3

To search categories having concept code c29867:

1. In the **Search** box, type c29867.
2. Select the **Category** check box.
3. Select the **Concept Code** option button.
4. Click the **Search** button.

The result of this search is as shown in the screenshot below.

The screenshot displays a web-based search interface. At the top, a search box contains the text 'c29867', followed by a 'Search' button. Below the search box is a hint text: '(E.g. Participant, Gene, Experiment, Protein, Specimen)'. A section titled 'Advanced Search' is expanded, showing several filter options: 'Category' (checked), 'Include Description' (unchecked), 'Attribute' (unchecked), 'Permissible Values' (unchecked), 'Text' (radio button), and 'Concept Code' (selected radio button). Below the filters, a blue banner indicates 'Search Results :- Total results (3)'. Three results are listed, each with a blue underlined link and a descriptive text: 1. 'CollectionProtocolRegistration (caTissue Core v1.1)' with the description 'A registration of a Participant to a Collection Protocol'. 2. 'Participant (caTissue Core v1.1)' with the description 'An individual from whom a specimen is collected'. 3. 'ParticipantMedicalIdentifier (caTissue Core v1.1)' with the description 'A medical record identification number that refers to a Participant'.

Figure 33: Search Result of Example 3

Concept code based search - Example 4

To search categories whose attribute has concept code C42614

1. In the **Search** box, type C42614.
2. Select the **Attribute** check box.
3. Select the **Concept Code** option button.
4. Click the **Search** button.

The result of this search is as shown in the screenshot below.

(E.g. Participant, Gene, Experiment, Protein, Specimen)

Advanced Search

☐ Category ☐ Include Description
☒ Attribute ☐ Permissible Values
☐ Text ☒ Concept Code

Search Results :- Total results (110)

[Gene Annotation](#)
 Gene annotation includes information about a gene such as its symbol, gene name(s), c....

[Literature-based Gene Association](#)
 Genes that are co-mentioned in the same PubMed abstracts have been curated. Each pai...

[Microarray Annotation](#)
 Microarray annotation includes information about a probe or probeset found on a micro....

[Orthologous Gene](#)
 Category providing details of orthologous genes associated with given gene obtained f...

[Array \(caArray v1.1\)](#)
 The physical substrate along with its features and their annotation

[ArrayDesign \(caArray v1.1\)](#)
 Describes the design of an gene expression layout. In some cases this might be virtu....

<< < 1 2 3 4 5 > >>

Figure 34: Search Result of Example 4

Define Limit: Operators and Values


This appendix explains the various operators and how to specify values while defining limits on a category.

The following table explains the different operators and provides examples as to how values are to be specified.

Table 3

| Operator | Data types for which applicable | Sample condition | Expected Results |
|------------------------|---------------------------------|---|--|
| Equals | Numeric Date String | Name equals "John" | Returns only John and not john. |
| Not equals | Numeric Date String | Identifier not equals 32 | Returns all records except those having identifier = 32 |
| Less than | Numeric Date | Participant's birth date less than 2000/1/1 | Returns all Participants born before January 1 st , 2000 |
| Less than or equal | Numeric Date | Signal value less than or equals to 0.5 | Returns all records for which signal value is not more than 0.5 |
| Greater than | Numeric Date | Cell count is greater than 10000 | Returns all records with Cell count > 10000 |
| Greater than or equals | Numeric Date | Specimen's collection-date greater than or equal to 2005/5/1 | Returns all specimens collected on or after May 1 st , 2005 |
| Between | Numeric String Date | Patient's admission date between 2000/1/1 and 2000/12/31 | Returns all Patients admitted in the year 2000 |
| Is null | Numeric String Date | Patient's contact number is null | Returns all patients who do not have any contact number |
| Is not null | Numeric Date | Participant' middle name is not null | Returns all Participants for whom middle name is not specified |
| In | Numeric String Enumerated | Signal value in 0.1,0.5,0.7 | Returns all records whose signal value is 0.1 or 0.5 or 0.7 |
| Not in | Numeric String Enumerated | Participant's gender not in male, unspecified | Returns all participants whose gender is either not specified or gender is not male. |

| | | | |
|-------------|--------|--|--|
| Starts with | String | Participant's first name starts with Jo | Returns all participants whose first name starts with Jo like John, Jona. Does not return the participant with name Scott, Henry, Jen |
| Ends with | String | Participant's last name ends with n | Returns all participants whose last name ends with n like John, Jon, Jen. Does not return participant with name Scott, Henry |
| Contains | String | Participant's first name contains ann | Returns all participants whose first name is Joanna, Anna, and Johanna. Does not return participant with name Jen, Denny |

- For attributes with enumerated values and Boolean attributes, clear a value by clicking on the value with the CTRL key pressed.
 - If you use the **In** operator for a non-enumerated attribute, you can specify multiple values as follows:
 - Multiple values should be separated using a comma. For example, to specify the constraint **Identifier IN 2, 4, 5, 7**, you should type the value as **2,4,5,7**.
- If a value itself contains a comma, you can enclose the value in double quotes. For example, to specify the constraint **Name IN "Einstein, Albert", "Newton, Isaac"** the value that you should enter is **"Einstein, Albert", "Newton, Isaac"**.
- For attributes that requires a date value, you can click the  icon to launch a date selector. You can also type a date string directly in the text box. The format of this string must be **YYYY/MM/DD**. A valid date string becomes **green** whereas an invalid date string turns **red**. An invalid date string will be ignored.

Record Details View

The details of a record usually consist of the values of the attributes of the category. See Figure 15: Results' Details for an example.

This section provides examples of special categories that result in a slightly different view of the record's details.

Administrator defined categories

These categories are built by grouping relevant attributes from one or more existing categories.

For each record of such a category the following information is shown:

Values of the attributes in a table

Related information as a tree. The leaf nodes of this tree are attributes; you can see the values of those in the **value** column.

Path: [Genomic Identifiers](#) >> **Genomic Identifiers_18667**

| Attribute | Value |
|-------------------------------------|-----------------|
| Ensembl Gene ID | ENSG00000120738 |
| Entrez Gene ID | 1958 |
| Uni Gene Cluster ID | Hs.326035 |
| Associated Information | |
| MessengerRNA (3) | |
| MessengerRNA_6174 | |
| MessengerRNA_6175 | |
| MessengerRNA_198617 | |
| GenBank Nucleotide Accession Number | |
| RefSeq mRNA Accession Number | NM_001964 |
| Ensembl Transcript ID | |
| Protein (5) | |
| Protein_372670 | |
| Ensembl Peptide ID | ENSP00000239938 |
| UniProtKB Primary Accession | P18146 |
| GenBank Protein Accession Number | |
| RefSeq Protein Accession Number | NP_001955 |
| Protein_372672 | |
| Protein_3435049 | |
| Protein_5676098 | |
| Protein_5852919 | |

Add To Data List

Apply Data List

Figure 35: One record of an administrator defined category (Genomic Identifiers)

BioDataCube

A biodatacube is a three-dimensional array representing the micro array data. The system displays a record containing a biodatacube (for example, the *BioAssayData* category of caArray) as shown in the screen below:

| Attribute | | | Value | | |
|------------|---------|--------------|--------------|-------------|---------------------|
| Identifier | | | 1 | | |
| Name | | | 1 | | |
| | | | | | |
| | Pairs 1 | Pairs Used 1 | Signal 1 | Detection 1 | Detection P-value 1 |
| 92555_at | 20.0 | Absent | 0.13876513 | 2.188886E-4 | 0.0030666848 |
| 92558_at | 16.0 | Present | 0.3276513 | 3.188886E-4 | 0.0030666848 |
| 92559_at | 20.0 | Present | 0.5645876513 | 4.188886E-4 | 0.0030666848 |
| 92568_at | 16.0 | Marginal | 0.464376513 | 5.188886E-4 | 0.0030666848 |
| 92574_at | 20.0 | Absent | 0.235876513 | 6.188886E-4 | 0.0030666848 |
| 92555_at | 20.0 | Absent | 0.13876513 | 2.188886E-4 | 0.0030666848 |
| 92558_at | 16.0 | Present | 0.3276513 | 3.188886E-4 | 0.0030666848 |
| 92559_at | 20.0 | Present | 0.5645876513 | 4.188886E-4 | 0.0030666848 |
| 92568_at | 16.0 | Marginal | 0.464376513 | 5.188886E-4 | 0.0030666848 |
| 92574_at | 20.0 | Absent | 0.235876513 | 6.188886E-4 | 0.0030666848 |
| 92555_at | 20.0 | Absent | 0.13876513 | 2.188886E-4 | 0.0030666848 |
| 92558_at | 16.0 | Present | 0.3276513 | 3.188886E-4 | 0.0030666848 |
| 92559_at | 20.0 | Present | 0.5645876513 | 4.188886E-4 | 0.0030666848 |
| 92568_at | 16.0 | Marginal | 0.464376513 | 5.188886E-4 | 0.0030666848 |
| 92574_at | 20.0 | Absent | 0.235876513 | 6.188886E-4 | 0.0030666848 |
| 92555_at | 20.0 | Absent | 0.13876513 | 2.188886E-4 | 0.0030666848 |
| 92558_at | 16.0 | Present | 0.3276513 | 3.188886E-4 | 0.0030666848 |

Figure 36: BioDataCube Record

This table displays the values of the attributes of the category, in a table at the top. A separate two-dimensional array shows the contents of the BioDataCube (the third dimension of the cube is transformed into multiple columns).

Types of Filters

A filter control is shown when you choose to filter the data of a data-category (see How to filter data).

The type of filter control upon the column based on which you wish to filter data.

Table 4: Column Type and Applicable Filter

| Data type of the column | Applicable filter |
|--------------------------------|-------------------|
| Numeric | Range filter |
| Text | Pattern filter |
| Column with permissible values | Enumerated filter |
| Boolean | Enumerated filter |

Using the Range Filter

The system shows a **Range Filter** if the column selected as a basis for the filter has numeric values.. It allows you to select the desired minimum and maximum values of the column.

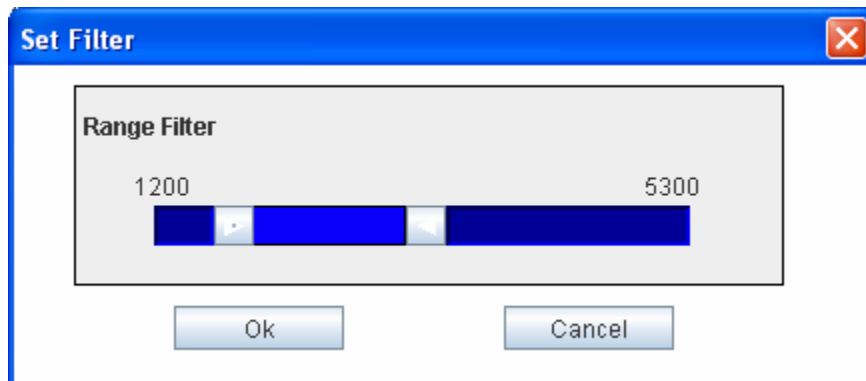


Figure 37: Set Filter for Range

There are two sliders, the left slider and the right slider. The value denoted by each slider is shown at the two ends.

1. Position the left slider at the minimum value desired.
2. Position the right slider at the maximum value desired.
3. Click **Ok** to apply the filter.

Using the Pattern Filter

The system shows a **Pattern Filter** If the column selected as a basis for the filter has text values. It allows you to specify a pattern that the values of that column should satisfy.

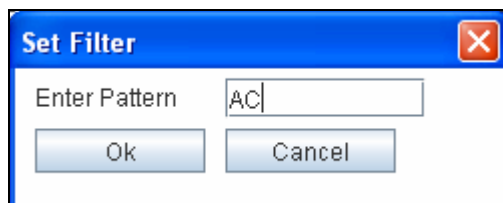


Figure 38: Pattern Filter

In the **Enter Pattern** field, type the desired pattern and click **Ok**.

Using the Enumerated Values Filter

The system shows an **Enumerated Values Filter** if the column selected as a basis for the filter has a set of permissible values. It allows you to specify the values that the column should satisfy.

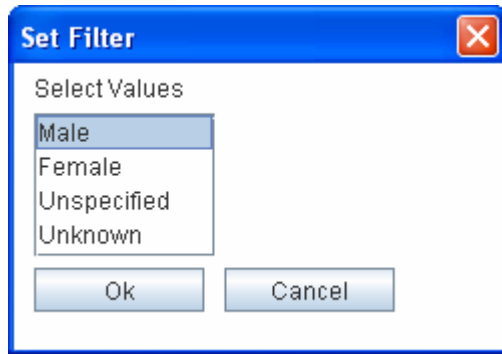


Figure 39: Enumerated Values Filter for the column gender

Select the desired permissible values and click **Ok**.